DOCUMENT RESUME

ED 029 901 24 TE 499 942

By-Owen, Curtis D., Jr.

A Study of Criteria for the Evaluation of Secondary School Instrumentalists When Auditioning for Festival Bands. Final Report.

East Texas State Univ., Commerce.

Spons Agency-Office of Education (DHEW), Washington, D.C. Bureau of Research.

Bureau No-BR-8-G-047

Pub Date May 69

Grant-OEG-7-9-000047-0089-010

Note-126p.

EDRS Price MF-$0.50 HC-$6.40

Descriptors-*Bands (Music), Evaluation, Evaluation Criteria, *Evaluation Methods, Evaluation Techniques, Instrumentation, Measurement Techniques, *Music, Music Education, *Performance Criteria, Performance Factors, Secondary Education, *Student Evaluation, Tape Recordings, Test Reliability

This study purposed to develop a means of evaluating student performance in auditions for festival bands which would minimize the inconsistencies of subjective judgment. Tape recordings of student auditions were played three times to judges. During the first audition, evaluators rated students on a continuum scale with numerical divisions, and during the second audition, they rated them on a scale containing verbal descriptions of the performance characteristics. In the third session, judges heard a professional instrumentalist play the audition music and then listened to and evaluated the student's performance using the verbal score sheet. Students were ranked by their scores, and interjudge reliability measured by analysis of the rank order was significant at the .01 level of confidence. Conclusions were (1) the three scoring techniques were equally reliable for evaluating musical performance, (2) the second technique was preferred by most adjudicators, (3) the performance characteristic of musicality was the most consistent factor in determining rank order, and (4) the graphic scale evaluation concept can be used effectively to score and rank instrumentalists who audition for festival bands. (Author/LH)

BR 8-G-047

PA 24

OE-BR

ED029901 TE

FINAL REPORT
Project No. 8-G-047
Grant No. OEG-7-9-000047-0089-(010)

A STUDY OF CRITERIA FOR THE EVALUATION OF SECONDARY SCHOOL INSTRUMENTALISTS WHEN AUDITIONING FOR FESTIVAL BANDS

Curtis D. Owen, Jr.
East Texas State University
Commerce, Texas 75428

May 1969

U.S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
Bureau of Research

TE 499 942

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

Final Report

Project No. 8-G-047

Grant No. OEG-7-9-000047-0089-(010)

A STUDY OF CRITERIA FOR THE EVALUATION OF SECONDARY SCHOOL INSTRUMENTALISTS WHEN AUDITIONING FOR FESTIVAL BANDS

Curtis D. Owen, Jr.

East Texas State University

Commerce, Texas

May, 1969

The research reported herein was performed pursuant to a grant with the Office of Education, U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

U.S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
Bureau of Research

ABSTRACT

A STUDY OF CRITERIA FOR THE EVALUATION OF SECONDARY SCHOOL INSTRUMENTALISTS WHEN AUDITIONING FOR FESTIVAL BANDS

Curtis Dawson Owen, Jr., Ph.D.
East Texas State University, 1969

Adviser: John F. Moss

Purpose of the Study:

The primary purpose of this study was to develop a means of evaluating student performance in auditions for festival bands which would minimize the inconsistencies of subjective judgment. Present audition conditions were studied in order to determine not only types of methods being used but also criteria for the construction of a scoring sheet which could be used under present organizational arrangements.

A secondary purpose of this study was to stimulate interest in further research in the area of musical performance evaluation.

Procedure:

The design of this study was of an experimental nature. The fact that no research report concerning audition techniques

was found necessitated a pilot study which was conducted to test procedural methods as well as tape recording and editing techniques. A survey of audition techniques being used in Texas all-region band auditions provided an opportunity for a comparison of different concepts.

Three audition techniques were developed for use in evaluating tape recordings of students recorded on tape in live auditions. Two scoring sheets were constructed for use in the audition sessions. They were based on numerical and verbal description concepts which have been used for rating many areas where subjective decisions must be made, i.e., merit rating by management, attitude scales, etc. In the third audition session a recording of the audition music by a professional instrumentalist was used to provide a standard of performance for adjudication.

Three adjudicators were used to evaluate the recorded performance of students who auditioned for all-region band in Region II and North zone of Region IV of the Texas Music Educators Association. Recordings of seventy-seven students were divided by instruments: twenty flutes, twenty clarinets, twenty cornet-trumpets, and seventeen trombones. Each adjudication panel evaluated the tape recordings of one instrument group three times using a different evaluation technique each time.

The scoring sheets were evaluated, and the score was used to place students in a rank order. Statistical analysis of the rank order provided an interjudge reliability coefficient of the three adjudicators and a correlation coefficient of each adjudicator's ranking with a pre-determined actual ranking. The scoring technique, which was thought to be most promising and which the adjudicators seemed to prefer, was re-evaluated with the recordings of twenty flute players.

Findings:

Statistical analysis of the rank order indicated that the interjudge reliability of all three audition techniques was significant at the .01 level of confidence. When the rank order of each adjudicator for each audition technique was compared with a pre-determined actual ranking for each instrument, all but one of the correlation coefficients were significant at the .01 level. The coefficient of the one exception was significant at the .05 level.

An analysis of six individual performance characteristics (tone, pitch, rhythm, technical accuracy, musicality, and sight reading) indicated that the evaluation of musicality produced the most reliable ranking. A survey of preferences of experienced adjudicators in the State of Texas indicated that they felt musicality should receive the largest weight in evaluation.

A re-evaluation of the second audition technique with a different adjudication panel produced an equally reliable ranking of student performances. Adjudicators were able to use the continuum concept very efficiently, and a minimum of time was necessary to instruct them in the proper way to mark their evaluations.

Conclusions:

1. The three scoring techniques developed in this research provide a reliable means of scoring musical performance evaluation.
2. The second technique was preferred by most of the adjudicators; and although not statistically more significant at the .01 level, the coefficients seemed to be a little higher.
3. The performance characteristic, musicality, was the most consistent factor in determining the rank order when compared to tone, pitch, rhythm, technical accuracy and sight reading.
4. The graphic evaluation scale concept can be used effectively to score and rank instrumentalists who audition for festival bands.

ACKNOWLEDGMENTS

A study which involves many individuals depends on the cooperation and encouragement of these individuals if the research is to be completed. I would like to express my appreciation to the high school band directors in every area of the State of Texas for their assistance in the development of practical ideas which could be used in the present structure of festival auditions. I would like to thank the band directors and applied music instructors in the four universities where the audition sessions were held for their interest and participation in this study.

In the most sincere way, I would like to thank Dr. John Moss for his direction of the study from the idea stage through the many facets of development and for his careful evaluation of the final report. To Dr. Neill Humfeld I would like to express my gratitude for his probing questions, his continuous encouragement, and his confidence as a teacher and friend. Special recognition must be given to Dr. James Richards for his creative ideas concerning auditioning techniques. I would like to indicate my deep appreciation to Dr. Grady Tice and Dr. Russell Davis for their instruction and counsel both in the classroom and during the course of the study.

Much of the strength and well-being which accompanied me while conducting this study was furnished by my wife,

Beverly. To her goes my deepest gratitude for the many hours she spent typing this report and my sincere thanks for her patience during the many stages of its development.

Curtis D. Owen, Jr.

TABLE OF CONTENTS

CHAPTER PAGE

I. INTRODUCTION, THE PROBLEM, AND DEFINITION OF TERMS 1

Introduction 1

The Problem 4

Statement of the problem 4

Purpose of the study 7

Significance of the study 8

Delimitations 8

Assumptions 9

Definition of Terms 10

Actual ranking 10

Adjudicating 10

Adjudicator 10

All-region band 10

Audition 11

Cornet-trumpet 11

Face validity 11

Festival band 11

Halo effect 11

Prepared exercise 11

Ranking 12

Rating 12

Recording technicians 12

CHAPTER PAGE

Scoring sheet 12

Sight reading 12

Value scale 12

II. LITERATURE RELATED TO PROBLEM 13

Development of Contests and Festivals in the Schools 13

Establishment of Festival Bands in Texas . . . 15

Adjudication 17

Summary 20

III. PROCEDURE AND DATA COLLECTION 22

Development of the Scoring Sheets 23

Pilot Study 31

Selection of Adjudicators 37

Population and Sample 40

Population 41

Sample 42

Recording and Editing Techniques 44

Recording techniques 44

Editing techniques 47

Data Collection 49

Preliminary steps to data collection 49

Data collection 52

Summary 55

CHAPTER PAGE

IV. ANALYSIS OF RESULTS 57
  Statistical Methods 57
  Comparison of Audition Data 59
    Interjudge reliability 59
    Rank-difference correlation 60
    Interjudge reliability of performance characteristics 67
V. SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS 75
    Summary 75
      General summary 75
      Adjudicator reaction 77
    Conclusions 78
      General Conclusions 78
    Recommendations 81
      Recommendations for application 81
      Recommendations for further research 84
SELECTED BIBLIOGRAPHY 86
APPENDICES 89
  Appendix A. Map of Regions and Areas 90
  Appendix B. Adjudication Scoring Sheets 92
  Appendix C. Survey Letter for Weighted Scale 95
  Appendix D. Text of Adjudicator Instructions 97
  Appendix E. Equal and Weighted Value Scales 102
  Appendix F. List of Rank Orders 105

LIST OF TABLES

TABLE PAGE

I. Random Sequence Used for the Order of Twenty Students Recorded on each of Four Instruments (Flute, Clarinet, Cornet-Trumpet, and Trombone) at All-region Band Auditions for Region II and Region IV of State of Texas, December, 1968 48

II. Kendall's Coefficient of Concordance W for Interjudge Reliability of Three Adjudicators 61

III. Spearman's Rank-Difference Correlation _rho_ of Adjudicators and Actual Rankings for Flutes 63

IV. Spearman's Rank-Difference Correlation _rho_ of Adjudicators and Actual Rankings for Clarinets 64

V. Spearman's Rank-Difference Correlation _rho_ of Adjudicators and Actual Rankings for Cornet-Trumpets 65

VI. Spearman's Rank-Difference Correlation _rho_ of Adjudicators and Actual Rankings for Trombones 66

LIST OF FIGURES

FIGURE PAGE

1. A Comparison of the W Coefficients for Tone. . . 68
2. A Comparison of the W Coefficients for Pitch . . 69
3. A Comparison of the W Coefficients for Rhythm 70
4. A Comparison of the W Coefficients for Technical Accuracy 71
5. A Comparison of the W Coefficients for Musicality 72
6. A Comparison of the W Coefficients for Sight Reading 73

CHAPTER I

INTRODUCTION, PROBLEM, AND DEFINITION OF TERMS

I. INTRODUCTION

Today in the secondary schools of our country, students are involved in many activities. Some of these activities are concerned with group participation, while others involve individual students. The high school band is one of these activities in which students perform as a group and also as individuals. Contests and festivals for these band students are provided by statewide interscholastic league organizations, state music teachers associations, colleges and universities, and many area, regional, district, and local groups.

Band contests and festivals have provided the bands with an opportunity for performance as well as a means of motivation toward excellence through competition with like organizations. Music directors were able to recognize the values which could be derived from contests and festivals early in the development of music competition. In a symposium on contests and festivals conducted during the 1936 convention of Music Educators National Conference, C. Stanton Belfour gave the following appraisal of the contribution the music contest makes to instrumental music organizations:

> Any survey of the American musical scene reveals the important role of music contest to the community.

> The joy of achievement is response to a challenge. A contest of groups is one kind of challenge to any community. It has been found from experience that the "state-wide community" offers the solution for the field of inter-school contests. Rural areas can best be grouped into the county unit to serve as the community for recognition of achievement. . . . From the ranking of players to first chairs in sections of bands and orchestras to the ranking [rating] of school units representing their respective communities--town, township, city, county, district, and state--the stimulus afforded by contests has brought Bach, Beethoven, Brahms, and the others of the glorious company of craftsmen to many communities which, without a contest program, would be denied opportunities to achieve.[1]

The motivation of an award for an outstanding performance was a strong factor in the development of high musical performance standards in our schools. Many methods of awards were tried in search of a way to praise the outstanding bands and to encourage the bands which were of lesser ability. In the early years, musical organizations were awarded positions in a rank order when they performed, i.e., first place, second place, etc. This was changed in the early 1930's to a rating system. Many rating systems were used and examined before the five-place rating plan was adopted in 1936 by the American Bandmasters Association, the National School Band Association, and the National School Orchestra Association. Some of the reasons for this change were given by Adam P. Lesinsky, who was at the time President of the National School Orchestra

---

[1] C. Stanton Belfour, "What Have Contests Done for Music Education?" Music Educators National Conference Yearbook (Chicago: Music Educators National Conference, 1936), p. 340.

Association. He stated:

> After ten years of participation in band, orchestra, solo, and ensemble contests, I still like contests. Many changes have taken place in this activity during these years. The bitter rivalry which existed between schools during the earlier days of contests has been transformed into a friendly competition. Directors, in the main, do not enter a contest any more with the idea of beating the other fellow. Their motive now is to have their work evaluated on the basis of perfection. The highest honor is no longer limited to one individual or group. Just as it is possible for more than one person in an algebra class to make an "A," so it is possible for more than one individual or group to make a first rating in the music contest, where the rating system of judging is used.[2]

The rating plan for evaluating the concert band, solos, and ensembles has proven to be successful and is still being used in most festival-contests today. However, students must still be ranked for seating in their own band or a festival band for which they are auditioning. Since only one student can be the first chair player in each section and since most band directors seat students in bands according to their performance ability, a ranking system must be used.

The recognition for progress in musical development is a strong motivating factor to the young instrumentalist. The student who auditions for a place in a festival band should be assured that the many hours of practice which must be spent in preparation for an audition will be carefully and accurately evaluated. In most cases the rankings given at an audition are carefully determined. The majority of the adjudicators who rank these students are honest and sincere in their evaluations.

---

[2] Adam P. Lesinsky, "The Festival-Contest," Music Educators National Conference Yearbook (Chicago: Music Educators National Conference, 1934), p. 276.

William E. Whybrew made the following remark in the conclusion of his book, Measurement and Evaluation in Music:

> Evaluation of musical performance probably never can be entirely objective due to its very nature, but systematic procedures can improve the reliability, and thus the validity, of such appraisals. It is to be hoped that clear definitions of standards and further research into procedures will help to effect still greater improvement.[3]

This research was designed to provide a study of auditioning procedures and an evaluation of audition techniques.

II. THE PROBLEM

Statement of the problem. Every year in the State of Texas alone, approximately 50,000 secondary school band students participate in auditions for festival bands. These include all-district bands, all-region bands, and many privately sponsored festival bands. In these festivals the students audition for a position in the bands, and are placed in the bands by rank according to the results of the auditions. In such auditions the students must depend on the subjective judgment of the adjudicators to evaluate and rank them.

---

[3] William E. Whybrew, Measurement and Evaluation in Music (Dubuque, Iowa: William C. Brown Company, 1962), p. 177.

Many methods have been tried in search of a fair evaluation of students' performances in auditions. For example, one method that has been employed has required two or more musicians to evaluate a student's performance in an audition, thus limiting the individual subjectivity of one adjudicator. This method has improved the reliability factor of the evaluation, but has not provided the desired result. Whybrew stated:

> The best tool for evaluating musical performances seems to be a kind of rating scale which will help auditors [adjudicators] to systematize their judgments and also aid in ensuring that all auditors [adjudicators] evaluate a performance on the same basis, i.e., give attention to the same elements of performance and weight the elements in the same way.[4]

Whybrew gave this analysis of the current methods of evaluating musical performance:

> Evaluation of performance is of necessity subjective. Too frequently, applied music exams and auditions are rated rather haphazardly through general impressions. More systematic procedures are needed in order to produce more reliable evaluations.[5]

The fact that a large number of students in secondary school bands audition each year for festival bands almost demands that music educators construct a method of evaluation which will effectively evaluate student performance. The subjective nature of these judgments makes this task difficult. The lack of research done in this important area is appalling.

---

[4] Ibid., p. 166.

[5] Ibid., p. 173.

Whybrew described the nature of the problem inherent in subjective evaluation in the following statement:

> Subjective ratings are notoriously unreliable. Indeed, the demonstrated unreliability of subjective grading added great impetus to the move toward objective measurement.[6]

The Watkins-Farnum Performance Scale[7] is an attempt to evaluate performance from an objective standpoint. Students are marked for pitch, time, change of time, expression, slurs, rests, holds, and pauses or repeats. The beauty of tone, to many musicians, the most important factor of all, is not part of the evaluative criteria.

> The Watkins-Farnum Performance Scale is the only objective test of performance currently available, and, as such, represents an important contribution to music testing. There is clearly a great deal that remains to be done in the measurement of musical performance.[8]

Even with its limitations, the Watkins-Farnum Performance Scale is a good method of evaluating student performance when time is not a factor. However, in auditions for festival bands, time is of the essence, since the average amount of time allowed for these auditions is between six and ten minutes. During this time the student must receive his instructions for the audition, play the required music, and the adjudicator must score the student's performance, record it on an evaluation sheet, and prepare for the next student.

---

[6] *Ibid.*, p. 68.

[7] John G. Watkins and Stephen E. Farnum, *Watkins-Farnum Performance Scale* (Winona, Minnesota: Hal Leonard Music, Inc., 1962).

[8] Paul R. Lehman, *Test and Measurements in Music* (Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1968), p. 71.

The problem, therefore, is the lack of a tested audition technique which can be applied to auditions for festival bands where limited time is a factor. This lack of a tested technique permits situations where auditions may be rather haphazardly evaluated through general impressions. These situations decrease the confidence which band directors and student instrumentalists place on the value of auditions for festival bands.

Purpose of the study. The primary purpose of this study was to develop a means of evaluating student performance in auditions for festival bands which would minimize the inconsistencies of subjective judgment. Present audition conditions were studied in order to determine not only the type of methods being used, but also the criteria for the construction of a scoring sheet which could be used under the present organizational arrangements.

A secondary purpose of this study was to stimulate interest in further research in the area of musical performance evaluation. It was the desire of the author that others may use this study as a basis for research into many related areas of adjudication.

Significance of the study. This study was significant because it tested auditioning procedures through controlled research conditions (described in Chapter III). In addition, scoring (evaluation) sheets which were developed by the author provided a clear, fast-scoring technique which could be used within the present structure of festival band auditions.[9] Although the scoring technique developed by this study was the first known application of this method to be tested through research, it represented a step toward providing a more reliable evaluation of individual musical performance. By focusing attention on auditions for festival bands, the study provided a guide to more effective methods of evaluation for the largest number of students who participated.

Delimitations. The study population was limited to students who entered and auditioned for all-region band on one of the following instruments: flute, clarinet, cornet-trumpet, and trombone, since these were the four largest groups of wind instruments who normally audition.

The regions involved in the research were Region II and the North zone of Region IV of the Texas Music Educators Association. (See Appendix A for a map of the regions and areas designated in Texas for music competition.)

---

[9] Ibid., p. 75.

The study was limited to twenty students each on flute, clarinet, and cornet-trumpet, selected at random from the order of performance determined by the region officers. In the case of trombones, only seventeen, from an original list of thirty-six, were present for the auditions. This was the total population at the audition.

The audition music each student played was selected by the region officers.

The students were scored only on the following six performance characteristics: tone, pitch, rhythm, technical accuracy, musicality, and sight reading.

Assumptions. The following assumptions were made for the study:

1. The scoring sheets used in the twelve auditioning sessions had "face" validity.

2. The "halo effect" caused by each judge hearing the same students three times was not significant due to the more than two-week time lapse between audition sessions and the changing of the order in which students were heard before each auditioning session.

3. Tape recordings, recorded with professional quality microphones and tape recorders, when played back on quality reproduction equipment, gave an accurate representation of the students' performances.

4. A musician who (a) has a professional teacher's certificate, (b) is presently directing a successful high school band program and (c) has been an adjudicator for all-region band auditions, can evaluate at least twenty students, recorded on tape, in rank order if adequate time is allowed for comparison and individual analysis of each student. This ranking is assumed to be the actual ranking in this study.

III. DEFINITION OF TERMS

These terms were given the following definitions when used in the study:

Actual ranking. The term, "actual ranking," is used to describe the rank order assigned to the tape recordings of the student instrumentalists by an expert adjudicator after a week of evaluation.

Adjudicating. Adjudicating is the process of evaluating musical performance in auditions.

Adjudicator. The adjudicator is a person who evaluates and scores the musical performances of students in auditions.

All-region band. An all-region band is composed of the students from a single region who are judged to have the highest level of performance skills in that region. The fifteen regions in Texas are determined by the University of Texas Interscholastic League. The regional festival is sponsored by the Texas Music Educators Association.

Audition. A trial musical performance to evaluate a band student's ability to play pre-determined musical examples is referred to as an audition.

Cornet-trumpet. In this study, the cornet and trumpet are referred to as one instrument, no distinction being made between the use of the cornet and trumpet since they are basically the same instrument as far as key, range, use in bands, and tone are concerned. The main difference is in their slight variation of physical structure.

Face validity. In this study, the term, "face validity," describes the fact that the scoring sheets look valid, particularly to those who are unsophisticated in test practices.

Festival band. A festival band is composed of students from several different bands, who assemble for a given period of time (usually from one day to two weeks) in order to play as a group. Participation is usually voluntary, and this band activity is usually promoted to motivate the development of and recognition of excellence in musical performance.

Halo effect. Halo effect is the influence a prior adjudicating experience has on the evaluator when he is faced with judging the same students.

Prepared exercise. The musical exercise which was assigned to the students to practice before the auditions is referred to as the prepared exercise.

Ranking. Ranking is achieved by listing the relative standing or position of the band students according to their ability as determined by their auditions, i.e., first place, second place, third place, etc.

Rating. Rating is the process of assigning a level of performance to an individual or group. This means that a Division I rating may be given for an outstanding performance to more than one person or group. On the other hand, for example, more than one band may receive a Division IV rating for a below-average performance. Most contests and festivals use a five-point rating plan for concert band, solo, and ensemble competition.

Recording technicians. The individuals who operated the recording equipment were skilled in the techniques of musical performance recording. They were the primary operators of the same recording equipment for the musical performances by the Music Department at East Texas State University.

Scoring sheet. A scoring sheet is the adjudication form which is marked by the person evaluating musical performance (adjudicator).

Sight reading. When the student plays a piece of music in the audition which he has never played before, this musical effort is called sight reading.

Value scale. The values assigned to the forty equal points along the continuum of the adjudication scoring sheet are described as the value scale.

CHAPTER II

LITERATURE RELATED TO PROBLEM

The information contained in this chapter is presented in order to cite significant writings of authorities in the area of auditioning. Some background of the festival-contest movement is included to show the nature of its development. A study directly related to the criteria of the evaluation in auditions was not found. The author examined the *Dissertation Abstracts*, the *Music Index*, the *Education Index*, the *Index of the Educational Research Information Center*, and all books and periodicals on music education, music tests, and the psychology of music in the East Texas State University Library. Additional study was made of related books and periodicals in libraries at North Texas State University and Baylor University. The purpose of this chapter is (a) to provide an organized presentation of events which created a need for this study, and (b) to give some background material into the nature of the problems of musical performance evaluation.

I. DEVELOPMENT OF CONTESTS AND FESTIVALS IN THE SCHOOLS

In the period from 1910 to 1925, the contest and festival movement took the initial steps leading to the annual events conducted today. The first competitions were primarily for bands and later for both orchestras and bands. Vocal contests followed a few years later. In 1936 the National School Vocal Association, the National School Band Association, and the National School Orchestra Association became auxiliaries of the Music Educators National Conference and assumed the leadership and coordination-control of National School Music Competition.[1]

From 1927 to about 1936, national contests were held annually in one location. The single national contest idea presented many problems such as a host city, transportation, absence time from school, and the tendency to localize the competition to states nearest the contest city. In 1937 a regional plan was proposed by the National School Band Association. By 1939 all three national organizations had approved a ten-region plan for national contests.[2]

Participation through regional contests in the United States grew from approximately thirty-one thousand in 1938 to about fifty-seven thousand in 1940.[3] This growth continued after World War II, and strong state organizations began to take the major responsibility for contests and festivals. In Texas, for example, the public school contests were placed under the joint direction of the Texas Music Educators Association and the University of Texas Interscholastic League in September 1946.[4] These two organizations have continued to organize and supervise these events for the schools in Texas until the present day. Since its beginning in 1946 the concert band contest governed by the University of Texas Interscholastic League has grown from approximately four hundred participants to over two hundred thousand.[5]

---

[1] "School Music Competition--Festivals," Music Educators National Conference Yearbook (Chicago: Music Educators National Conference, 1940), p. 508

[2] "Historical Sketch--Competition-Festivals," Music Educators National Conference Yearbook (Chicago: Music Educators National Conference, 1938), p. 430.

[3] Ibid.

II. ESTABLISHMENT OF FESTIVAL BANDS IN TEXAS

The first all-state band in Texas was assembled in Fort Worth in 1938. Two bands were selected from "high school musicians throughout the state."[6] The purpose of the all-state band was to play music which appeared on the national contest list as set forth by the National School Band Association. The guest conductors made comments to the band directors, who were present, about some of the performance problems inherent in each selection.

---

[4] Roy L. Higgins, "An Analysis of Adjudication Procedures for A-AA Bands in University of Texas League Competition " (unpublished Master's thesis, The University of Texas, Austin, 1965), p. 8.

[5] *Ibid.*, p. 10.

[6] Nelson G. Patrick (ed.), "Minutes of 1938 Annual Meeting," *Minutes and Procedures of Texas Music Educators Association* (Austin, Texas: Texas Music Educators Association, 1961), p. 221.

From 1938 to 1945 the Texas Music Educators Association was the organizer and sponsor of most state band contests including the concert band, marching band, solo and ensemble contests. These contests were discontinued during World War II.[7] When the University of Texas Interscholastic League and the Texas Music Educators Association merged their contest activities in 1946, the University of Texas Interscholastic League was given the task of contest supervision and regulation of the above-mentioned contests. The Texas Music Educators Association continued to have direction of activities of the clinic bands at its annual convention.

For the past twenty years the all-state bands and clinic bands at the annual Texas Music Educators Association Convention have obtained their student participants from the best players in the all-region bands of the State. These region bands have been organized by the band directors of each region who have been active members of the Texas Music Educators Association. In recent years some regions have found it necessary to have more than one all-region band with some type of regional elimination auditions for students who go to the all-state band. Participation in these regional contests has increased to the point that five area elimination auditions must be held in order to select only the most outstanding performers from each area for membership in the all-state clinic band. This increased interest in the festival band has grown to include all-district bands, all-city bands, summer band camp bands, all-senior honor bands, and many other types of performing organizations. These festival bands are not sponsored by the Texas Music Educators Association but are organized to provide band students with the opportunity to perform in a band which is composed of the best performers as determined by the auditions.

---

[7] Higgins, loc. cit.

III. ADJUDICATION

The process of placing a value measurement on the evaluation of a musical performance has proven to be very difficult. It has not been uncommon for one adjudicator to hear and evaluate between twenty-five and thirty-five students at one day's audition. After hearing as many as ten students, it becomes difficult to remember how well the second or third student played.

Peter W. Dykema gave the following observations of the art of judging in a report to the Music Supervisors National Conference in 1926:

> He who tries to judge must constantly compare what he hears with two standards; first, the absolute, and second, the relative. Absolute standards are the results of many preceding relative standards. One must have heard much music well performed, and one must be able to interpret with abundant imagination the suggestions of the scores. From this one obtains a conception of that

> perfection which serves as the absolute standard which is practically never embodied in a high school contest. It is used to establish balance, keep the blood cool and the mind clear when the long hours suppressed excitement, the tension of competition, and poor ventilation, all press upon the aching head of the adjudicator. The relative standard, for comparison, is that fine rendering which is the joy of every adjudicator's ears and heart. Nothing is pleasanter than to hear early in the contest a fine performance which may serve as a standard for marking those which precede and follow it. Nothing is more trying, nothing comes so near developing panic as to have a contest through six, eight, ten, twelve, fourteen numbers without a single one of them having been good enough to set up as a relative standard for this group.[8]

These thoughts of Dykema were made early in the band contest movement and may lack the insight and experience of present day adjudicators. However, even present day writers indicate that each adjudicator's major problem is the lack of a standard for evaluating performance. Neilson, writing on the basis of judgment in adjudication, said:

> Critical judgment is a balance determined by many factors. Some can be measured, others cannot. These include temperament, knowledge, intellect, experience, and the ability to maintain an objective relationship toward music and its performance. The rating given by an adjudicator is a result of these complexities. The fact is that there is often little if any resemblance from one person to another within the area and scope of the subjective complexities. Small wonder that every adjudicator has his own standards of marking and grading.[9]

Whybrew had the following analysis of the evaluation of musical performance:

---

[8] Peter W. Dykema, "Observations of an Adjudicator," Music Supervisors National Conference (Chapel Hill, North Carolina: Music Supervisors National Conference, 1926), p.99.

[9] James Neilson, "Role of the Adjudicator," The Instrumentalist, X (May, 1958), p. 22.

> Evaluation through general impression is likely to be highly unreliable. . . even when performed by the most careful person of unimpeachable musical qualifications.
>
> Musical performance, because of its very nature, is extremely difficult to evaluate reliably. Not only is it a highly complex affair but certain aspects of it have so far defied precise definition, to say nothing of precise measurement. . . . Further complicating the reliable evaluation of musical performance is the lack of definite standards of quality for various elements of performance.[10]

Whybrew suggested the use of a rating scale using a point system as the best method for evaluating musical performance. He concluded his discussion of rating scales with the following comment:

> Even with such guides, subjective evaluations will continue to be rather unreliable, but some such sort of analysis in rating seems far preferable to the general impressions which too frequently serve as the basis for evaluation.[11]

All three of these writers indicated that the art of adjudication depends on the ability of an adjudicator to place an accurate measurement on the intangibility of a musical performance. To the author, one of the main problems in adjudicating seemed to be the lack of a tested method for assigning a numerical or letter value to the evaluation of a musical performance. Whybrew suggested that a descriptive continuum might aid in scoring auditions. He said, however, that he doubted that this method would prove to be any more accurate.[12] The need for studying this problem in greater detail and testing different scoring techniques seemed to almost demand that someone try to improve the situation.

---

[10] Whybrew, op. cit., p. 165.

[11] Ibid., p. 172.

[12] Ibid., p. 171.

IV. SUMMARY

The music festival-contest movement in the United States has grown from a small community-wide activity initially to a large, highly organized and structured series of annual events. One part of these events has been the festival band movement. The lack of a tested means of measuring or scoring the evaluation of musical performance for festival bands has been the biggest problem faced by the organizers and planners of festival bands. Many festival band organizers feel that the best solution to this problem is to have the best musician available to adjudicate the auditions.

Dykema and Neilson have served as adjudicators in many parts of the country and have studied the problems inherent in the evaluation of musical performance. They have exposed one of the difficulties of musical performance evaluation as being the need to establish a standard of performance for the many groups of performers an adjudicator will hear. Whybrew indicated that some type of systematic procedure of evaluation might improve the reliability of this measurement.

The review of literature by the author has revealed that the writings in this area have many characteristics of the problem itself. The ideas of writers in the field have been based on general impressions which have been tempered with many years of experience. These subjective judgments are valuable as means of gaining insight into the problem but have little effect in the progress toward a solution of it.

CHAPTER III

PROCEDURE AND DATA COLLECTION

Due to the large number of public school students who audition for festival bands each year, there is a considerable need for a more reliable method of evaluating musical performance. A paucity of significant research concerning the evaluation of musical performance in auditions for festival bands, however, dictated that the procedural techniques used in this study had to be developed by the author.

This chapter describes the steps taken to develop procedural factors necessary and prerequisite to undertaking research in this area. The steps in the order of their presentation were:

1. Development of the scoring sheets
2. Pilot study
3. Selection of adjudicators
4. Population and sample
5. Recording and editing techniques

Only after the preceding preliminaries had been carefully developed could the actual data collection process begin. The data collection process provided the evaluation necessary to determine the degree of consistency made possible by the scoring techniques which were developed for this research.

I. DEVELOPMENT OF THE SCORING SHEETS

The study of criteria for the evaluation of musical performance began with an investigation of methods currently being used in festival band auditions. Since the all-region band organization in the State of Texas was the largest single group of organized festivals in the State, this group furnished a good starting point for the research. Current methods of adjudication in the fifteen regions of the State were surveyed. The results of this survey revealed that all the regions prescribed a percentage range for the adjudicator to use in evaluating certain phases of performance. The adjudicator would ascribe a score within the allowed percentage range which he felt represented the level of the audition performance.

The basic concept of a percentage range for the scoring of student performance was used in different forms in all regions. Over one-half the regions required the playing of major, minor, and chromatic scales for auditions in addition to prepared exercises and sight reading. The regions which required scales, prepared exercises, and sight reading divided the percentage range as follows: scales, 25 per cent; prepared exercises, 50 per cent; sight reading, 25 per cent; or: scales, 33 per cent; prepared exercises, 33 per cent; sight reading, 33 per cent. Regions which did not require scales ascribed the following percentage ranges to the elements of prepared exercises and sight reading: prepared exercises, 75 per cent;

sight reading, 25 per cent; or: prepared exercises, 80 per cent; sight reading, 20 per cent; or: prepared exercises, 50 per cent; sight reading, 50 per cent. Some indecision was noted concerning the proper weight which should be given to these divisions. Many of the regions indicated that their present weighting of these divisions had been recently adopted and, as yet, had not been evaluated for effectiveness.

A study of scoring sheets used by the regions for their instrumental auditions indicated that several regions had divided the percentage range allowed for the scales and prepared exercises. For example, one region divided the one hundred points as follows: two major scales, ten points; two minor scales, ten points; chromatic scale, five points; area audition music, twenty points; concert music, ten points; tone quality, ten points; intonation, five points; general technique, five points; sight reading, twenty-five points. Another region which did not use scales made the following division: prepared exercises, (a) technique (articulation, style, range, etc.) 40 per cent, (b) tone (intonation, quality, musicality, etc.) 40 per cent; sight reading, 20 per cent.

The above information revealed that the adjudicator must transfer his subjective evaluation to a point in a range of numbers in order to arrive at the student's score. Although the music requirements for the auditions indicated some variation and the weighting of each performance division provided some difference from region to region, the scoring concept remained generally the same for all regions.

The scoring sheet used at the Blue and Gold Band Camp in June 1968 at East Texas State University as part of the pilot study provided an opportunity for further study concerning evaluation criteria. For this instance, the number of performance characteristics was limited to three: rhythm, tone, and pitch. Sight reading was listed as a separate category. The scoring portion of the audition sheet was constructed as it appears below.

(CIRCLE ONE)

RHYTHM	1	2	3	4	5	6	7	8	9	10
TONE	1	2	3	4	5	6	7	8	9	10
PITCH	1	2	3	4	5	6	7	8	9	10
SIGHT READING	1	2	3	4	5	6	7	8	9	10

TOTAL_______

After the completion of the Band Camp auditions, the adjudicators indicated that three divisions of performance characteristics were not enough to properly evaluate audition performances. The concensus of opinion of the adjudicators was that two additional characteristics should be added to the ones which were used in the Band Camp auditions. These two additional characteristics would be technical accuracy and musicality. Technical accuracy was considered to be less subject to extreme interpretation than the word, "technique." Musicality was added for two reasons: (a) The adjudicators who participated in the Band Camp auditions felt that an evaluation of musicianship was essential, and (b) Some writers have expressed the idea that evaluation of musical ability should not be broken into parts but judged only as a whole. In reference to the second reason for including the characteristic of musicality, the author felt that any scoring sheet should reflect both sides of a controversy concerning evaluation of musical ability. The two sides of the controversy were debated by psychologists Carl E. Seashore and James L. Mursell. Seashore used the factors of pitch, loudness, duration, rhythm, timbre, and tonal memory as elements of sound which the mind can receive through the transmission of this sound by the ear. Mursell expressed his belief that musical talent was more than a set of specific attributes dependent on sensory capacities. Both psychologists conducted studies on which they based their ideas.[1] Although these theories were presented in the area of musical aptitude tests, they should be considered in connection with the performance characteristics which could and should be included in the evaluation of musical performance.

---

[1] William E. Whybrew, Measurement and Evaluation in Music (Dubuque, Iowa: William C. Brown, Company, 1962), p. 90.

Whybrew suggested an idea which business management and test and measurement specialists have used in rating scales requiring subjective evaluation. The evaluation sheet consisted of a verbal description which was placed directly above a continuum. Although Whybrew indicated that other areas of testing preferred the verbal descriptions to numbers, he was doubtful if verbal descriptions would be as reliable as numbers.[2] A search for different types of musical performance evaluation sheets failed to reveal that this method of evaluation had ever been used in instrumental auditions.

Guilford described this type of rating scale as a "graphic scale." He provided the following evaluation of the graphic rating scale:

> The virtues of graphic rating scales are many; their faults are relatively few. Among the advantages frequently cited in their favor are the facts that they are simple and easily administered; they are interesting and require little added motivation; they are quickly filled out; and they do not require the rater to bother with numbers. These features the rater finds attractive. From the point of view of the investigator, the graphic scale provides opportunity for as fine discrimination as that of which the rater is capable and the fineness of scoring can be as great as desired. As for disadvantages, there are none that do not apply to most other types of scales, except for somewhat greater labor of scoring in connection with some formats.[3]

---

[2] Ibid., p. 171.

[3] J. P. Guilford, Psychometric Methods (New York: McGraw-Hill Book Company, 1954), p. 268.

The continuum provided a means of evaluation which did not require the adjudicator to transcribe his subjective evaluation into a number. By picturing the best possible performance at the right end of the continuum and the worst possible performance at the left end of the continuum, the adjudicator became only an evaluator. He was not required to be a manipulator of numbers at the moment the rank order must be decided.

For this research project, two types of scoring sheets were developed using the continuum as the basis for scoring the evaluation. Numbers were placed above the continuum on the first scoring sheet, and verbal descriptions replaced the numbers on the second scoring sheet. (See Appendix B.) Descriptions were written for each of the six performance characteristics reported earlier. These descriptions reflected the best possible performance on the right and the poorest possible performance on the left.

Four of the verbal descriptions closely parallel the ones used by Whybrew in his example of pitch.[4] The characteristics of musicality and sight reading seemed to be best expressed with one-word descriptions. These descriptions were refined by two processes: (a) Five experienced teachers

---

[4] Written permission was obtained from William E. Whybrew to use this example.

"PERMISSION TO REPRODUCE THIS
COPYRIGHTED MATERIAL HAS BEEN GRANTED
BY William E. Whybrew
TO ERIC AND ORGANIZATIONS OPERATING
UNDER AGREEMENTS WITH THE U.S. OFFICE OF
EDUCATION. FURTHER REPRODUCTION OUTSIDE
THE ERIC SYSTEM REQUIRES PERMISSION OF
THE COPYRIGHT OWNER."

and adjudicators were asked to read, evaluate, and make constructive suggestions, and (b) The adjudicators who participated in the pilot study auditions made suggestions which they felt eliminated ambiguities in the wording of the descriptions.

The exact length of the continuum was determined in order that a scoring scale could be developed. In the pilot study, each of the six performance characteristics was ascribed a number to correspond to the point on the continuum where the adjudicator had placed a mark. The continuum was divided into forty equal parts. Each of the forty divisions received the accumulative value which corresponded to that point on the scale, i.e., the score which fell on the twenty-eighth mark from the left would receive the value of each division multiplied by twenty-eight. To facilitate scoring efficiency, the scoring scale was copied on a clear transparency, and the transparency was placed over the completed scoring sheet. This technique proved to be satisfactory and was used for scoring the evaluations of the actual research.

The scoring scale concept described above was used in two different ways. The first scoring scale was an equal value scale. On this scale, each of the performance characteristics was given the same value. The second was a weighted value scale. This weighted value scale was used because the survey of all-region scoring techniques seemed to indicate that the main variation between regions was the degree of weighting the different divisions were given. The author prepared a survey letter (See Appendix C.) which was sent to forty musicians in the State of Texas who were currently either directing a performing musical organization or supervising a public school music program. The letter described briefly the nature of the research and requested the individual to indicate the appropriate weight he felt each characteristic should receive. Geographically, the sample included every region in the State. The sample also included band directors of every band classification from junior high school through the largest university. A 70 per cent return indicated the following weights: tone, 21 per cent; pitch, 11 per cent; rhythm, 12 per cent; technical accuracy, 13 per cent; musicality, 24 per cent; sight reading, 19 per cent.

In summary, the scoring sheets were developed following: (a) a study of current audition scoring methods, (b) experimentation and evaluation of scoring sheets during a pilot study at the East Texas State University-sponsored Blue and Gold Band Camp, (c) an examination of literature dealing with scoring techniques employed in individual performance auditions, and (d) refinement of the scoring sheet as a result of the pilot study. The weighted scale for scoring was a result of the opinions of experienced band directors and supervisors. The primary significance of the scoring sheets which were developed was that the adjudicator could place a value on a performance without the limitation of a specific number.

II. PILOT STUDY

The purpose of the pilot study was (a) to test and evaluate recording techniques, (b) to experiment with tape editing methods to insure a minimum of quality loss, (c) to rehearse the procedural steps required to administer the audition of the tapes, (d) to determine the problems involved in scoring the continuum, and (e) to evaluate the results of statistical formulas as they applied to the analysis of the data.

The initial step in the pilot study was conducted at the Blue and Gold Summer Band Camp at East Texas State University in June 1968. The students who enrolled in the Band Camp were sent a copy of the audition music by the Band Camp secretary. The enrollment deadline was two weeks prior to the camp's opening day. Applications which were received on the deadline date permitted the secretary to mail the audition music to these students at least ten days before the audition. A list of major scales to be memorized and a statement that sight reading would be part of the audition was printed on the page directly above the prepared exercise.

Tape recordings were made during the auditions of twenty-three students on each of the following instruments: flute, clarinet, cornet-trumpet, and trombone. (No distinction was made between the use of the cornet or trumpet in these auditions.) Due to the large number of students who auditioned on each of the four instruments, it was necessary to divide the students and conduct auditions in two rooms for each instrument. The rooms were closely matched as far as size and acoustics were concerned, and tape recorders and microphones used to record each instrument were the same quality, i.e., two recorders of the same quality were required for each instrument. Four different types of recorders were used for the four instruments recorded. These were: (a) flutes, semi-professional portable stereo tape recorders with cardioid dynamic microphones; (b) clarinets, professional tape recorders and professional microphones; (c) cornet-trumpets, solid-state "stereo compact" portable tape recorders and dynamic microphones; (d) trombones, small monaural tape recorders and dynamic microphones.

The tape recordings revealed several important points concerning sound reproduction. First, monaural and stereo portable recorders, basically intended for use in the home or for recording the voice, were not of acceptable quality for this study. Secondly, semi-professional (This refers to an intermediate quality recorder which is usually priced between $350 and $500.) portable stereo recorders could produce an acceptable quality recording. Further testing revealed that if a professional quality microphone was used with semi-professional recorders, the quality of recording would be much better than it would be with cardioid dynamic microphones which were used during the pilot study. Thirdly, professional recorders and professional microphones provided the best quality recording. Professional tape recorders were very sensitive and required careful adjustment by an electronic technician who specialized in alignment and maintenance of high quality recording equipment. Only one-half of the clarinet recordings could be used for evaluation because moisture in one of the microphones caused some distortion.

The experience gained from these recordings provided answers to many questions concerning technical problems involved. Some conclusions which were made from this part of the pilot study were: (a) If professional recorders were used, they should be aligned and adjusted to original specifications, and professional microphones should be allowed to warm up for at least one hour before the auditions began in order to allow any moisture to evaporate from the heat of the electrical current. (b) If professional microphones were used with semi-professional portable recorders, the tape recordings would be of acceptable quality.

Methods of editing tapes were tested in several different ways. First, one set of tapes was edited by re-recording on another tape the performance of the student's audition. All conversation and lapses of time were removed from the tape. Loss of quality using this method to edit tapes was too great and, therefore, could not be used. The second method of editing tapes was to cut all conversation and time lapses from the original tape and splice together only performances of students. The original quality of sound was maintained; however, the change between prepared exercises was abrupt and did not allow a clear division of each exercise and the sight reading. This also made change of random sequence difficult since a splice was made for each of the four times a different sequence was used.

The methods of editing which provided the needed qualities were: (a) Cut from the original tape all conversation and time lapses and put a two-foot strip of blank tape between each prepared exercise and between the last prepared exercise and sight reading. (b) A three-foot strip of paper leader tape was spliced between each student's recording. The leader tape was labeled with the original number of the student's order in the auditions. This three-foot strip of white leader tape provided the length of tape necessary to splice the tape each time the random sequence was changed.

A few changes were made in the wording of descriptions on the scoring sheet as a result of the pilot study. Descriptions for tone, pitch, rhythm, and technical accuracy were not changed. One-word descriptions of musicality were changed from poor, fair, good, pleasing, and beautiful to poor, fair, average, good, and excellent. One-word descriptions of sight reading were changed from poor, fair, good, excellent, and flawless to poor, fair, average, good, and excellent. Adjudicators who scored the recordings of students at the Blue and Gold Band Camp were pleased with the ease of scoring which the continuum afforded.

Numerical scores of each performance characteristic were recorded and scored with a value scale which was reproduced on a transparency. The transparency was placed over the continuum, and the score was determined at the point where the adjudicator had placed a mark across the continuum. If the mark of the adjudicator fell exactly between two points on the value scale, the larger score was awarded. The total score of the six performance characteristics was summed for each student, and a rank order was determined.

The two statistical formulas which were applied to the ranks for analysis were Kendall's Coefficient of Concordance W and Spearman's Rank-Difference Correlation rho. Interjudge reliability was determined by using Kendall's Coefficient of Concordance W. An additional step was needed before the Spearman rho formula could be applied. In order to establish a constant factor, tapes of the pilot study were ranked by an experienced adjudicator who was given enough time to hear and compare students on the tape several times. The ranking which the adjudicator gave the student performance on the tape was assumed to approach the actual rank order. Results of the Spearman rho formula provided a coefficient of correlation between the ranking from the scoring sheets and the constant factor which was designated as the actual ranking.

Tape recordings of the flutes were evaluated three times by a panel of three adjudicators. In the first audition session the scoring sheets with numbers over the continuum were used, and in the second session verbal descriptions were placed over the continuum. The variable in the third audition session was a recording of a professional musician playing the prepared exercises which was played before each student's performance was heard. The recording was an effort to determine if a constant standard of performance might increase the consistency of the ranking.

In summary, the pilot study provided needed information concerning mechanics of operation, procedural techniques, and a real concept of the nature of administrating such a research project. An analysis of the statistical data indicated that an interjudge reliability, from Kendall's W, of over .80 could probably be expected. Results of the analysis of rho indicated that the difference between the consistency of the ranking of the scoring sheet and the actual ranking could be determined. Another significant contribution of the pilot study was the contribution it made in refinement of scoring sheets to be used in the research.

III. SELECTION OF ADJUDICATORS

Selection of adjudicators was considered to be one of the most important aspects of this research. Although writers in music education have had very little to say about methods of scoring subjective evaluation, much has been written about selection of the adjudicator. McAllister said in 1936: "No contest can really serve the purpose for which it was intended without satisfactory adjudicators. . ."[5] Patrick in his opening statements on qualifications for judges expressed this belief: "The values derived from the contest are in proportion to the competency of the judge and his ability to make critical evaluations."[6] Whybrew summed up the importance of qualified adjudicators with the following statement:

> No matter what other arguments may be marshalled in support of contests and festivals, it seems clear that only thru the sincere efforts of competent and conscientious adjudicators can any real value be realized.[7]

Wilcox in commenting about the qualities of the adjudicator said: "The judge must be a competent musician, familiar with public school problems, with a thorough understanding of what contests are intended to accomplish."[8]

---

[5] A. R. McAllister, "Contest Management," Music Educators National Conference Yearbook (Chicago: Music Educators National Conference, 1936), p. 353.

[6] Nelson G. Patrick, A Handbook for Adjudicators (Austin, Texas: The University of Texas Interscholastic League), p.5. (undated)

[7] William E. Whybrew, "The Adjudicator as Educator," The Instrumentalist, XIII (May, 1959), p. 91.

After stating that the adjudicator should have "an established reputation, a broad musical background. . . a sensitive ear and mature, discriminating musical taste," Bachman made the following analogy about qualities of the competent adjudicator:

> Not every fine musician will be a good judge. He should have what lawyers refer to as "judicial temperament." Many brilliant lawyers would fail as judges because of the lack of this quality, and some of the most distinguished judges have not always been superior lawyers. Some sensitive musicians fail to be satisfactory judges, perhaps because they are too sensitive. They may be so influenced by certain factors, of which they disapprove, that they fail to give proper weight to other more important qualities.[9]

The adjudicators selected to participate in this research were divided into three broad classifications: (a) high school band directors, (b) college band directors, and (c) solo instrumental performers. The last classification included college teachers of a specific applied music instrument. All these adjudicators had experience teaching instrumentalists in public schools. The range of years of experience was from five to over thirty years, and all were interested in the development of a more reliable evaluation method.

---

[8] E. W. Wilcox, "The Long View in Music Contests," Music Supervisors National Conference (Chapel Hill, North Carolina: Music Supervisors National Conference, 1927), p. 234.

[9] Harold B. Bachman, "Competent Adjudication" The Instrumentalist, XVII (March, 1963), p. 50.

Three adjudicators were used for each instrument. This number has been considered to be desirable since early in the contest and festival movement. Lesinsky recalled an occasion when he was a member of a three-judge panel to adjudicate solos and ensembles, ". . . where I judged a contest with two other judges we were asked how it was possible for one B-flat clarinet trio out of five entries to receive a first, third, and fifth place."[10] McAllister was of the opinion that "three adjudicators are preferable. . ." to one or two when possible.[11]

The type of adjudicating panel used in this study included the three classifications of adjudicators previously mentioned (one high school director, one college director, and one solo instrumental performer) and also fulfilled the three-adjudicator qualification which has been the preferred practice in contests and festivals of today. Since college directors provided the smallest population from which a selection could be made, they were selected first. College directors were: Neill Humfeld, East Texas State University; James Jacobsen, Texas Christian University; Donald Moore, Baylor University; and William Postlethwaite, University of Texas at Arlington. Applied instrumental instructors at these universities were contacted, and three of four men contacted were able to participate at the designated time. One instructor had a class schedule conflict and was unable to be present. The three applied instructors who could participate were: James Deaton, East Texas State University; James Mahoney, University of Texas at Arlington; and Bernard Smith, Baylor University. Mack Guderion, trumpet player in the Dallas Symphony Orchestra, was able to come to Texas Christian University at the time of auditions, and this addition completed the four men needed in the solo instrumental performer classification. Four high school directors from high schools in the area of the four colleges were selected to complete the panel. These men were: G. A. Carson, Arlington Heights High School, Fort Worth, Texas; Bob Cartwright, Greenville High School, Greenville, Texas; Don Filgo, Richfield High School, Waco, Texas; and Roger Winslow, L. D. Bell High School, Hurst, Texas. All these men accepted the invitation and adapted their busy schedules in order to be present to evaluate the audition tapes.

---

[10] Adam P. Lesinsky, "The Festival-Contest," Music Educators National Conference Yearbook (Chicago: Music Educators National Conference, 1934), p. 276.

[11] McAllister, loc. cit.

IV. POPULATION AND SAMPLE

In the State of Texas, the largest organized festival band is the all-region band, which is sponsored by members of the Texas Music Educators Association in each of the fifteen regions. One of the strong motivating factors in region band success is the fact that student instrumentalists, who are judged to be the best performers in each section, are eligible for the all-state band. The all-state band is organized by the Texas Music Educators Association to serve as a clinic band at the annual convention held in February of each year.

> TMEA coordinates all the details in forming these organizations [the all-state band, orchestra, and chorus], which represent the "cream" of over 500,000 youngsters in school music programs in Texas. TMEA involvement includes everything from selection of music, rigorous tryouts and auditions on the local and regional level, and the selection of nationally prominent conductors to rehearse and lead the groups in their final Grand Concert--traditionally the climax of each year's Clinic-Convention.[12]

A conservative estimate of the number of students who audition for all-region bands each year in Texas is approximately eighteen thousand.

Population. The population used for this study was composed of students who auditioned on flute, clarinet, and cornet-trumpet in the North zone of Region IV of the Texas Music Educators Association, and students who auditioned on trombone in Region II of the Texas Music Educators Association. The flute, clarinet, cornet-trumpet, and trombone sections provided the largest number of participants. In other words, more students auditioned on these instruments than on any of the other instruments. This was true because these were the largest sections in the majority of high school bands.

---

[12] "TMEA, What Is It Doing for Music Education in Texas?" (Houston, Texas: Texas Music Educators Association, 1968), p. 4.

Students who auditioned for all-region band were members of high school bands in their region and, in most cases, were screened by their band directors before being entered in all-region band auditions. In the two regions where the sample was taken, the range of size of schools and bands was from small high schools with less than one hundred students enrolled to high schools with an enrollment of over two thousand. The area of the two regions included thirty counties of the North and Northeast Texas area.

Sample. A sample of twenty students from each section--flutes, clarinets, cornet-trumpets, trombones--were recorded at random from all students on these instruments who auditioned the day of all-region band tryouts. The author obtained the number of entries on each instrument from the all-region band organizer. The number of entries were: thirty-eight flutes, forty-eight clarinets, sixty-two cornet-trumpets, and thirty-six trombones.

The random order was obtained for recording the flutes by placing thirty-eight slips of paper with the corresponding number written on the face of each slip into a large basket and by drawing a number from the basket after the basket had been shaken each time. After a number was drawn from the basket and recorded on paper, it was placed back in the basket and the basket shaken again. When a number which had been previously used was drawn. it was not recorded but put back into the basket. The random order for recording clarinets, cornet-trumpets, and trombones was obtained in the same manner as were flutes except that the number of slips of paper placed in the basket was different for each group. The drawing was repeated for each instrument until twenty-three different numbers had been drawn each time making a total of ninety-two drawings--twenty-three for flute, twenty-three for clarinet, twenty-three for cornet-trumpet, and twenty-three for trombone. Although the design of the study required only twenty students on each instrument, numbers for twenty-three students were drawn in case technical difficulties eliminated any of the original twenty recordings.

The list of the random sequence was used by technicians operating recorders during each audition. The number of students who actually auditioned was smaller than the number who officially entered. In the case of flutes and trombones, the number of students who auditioned was twenty-three and seventeen respectively. Because of this small number, all these students were recorded. Of forty students on clarinet and forty-nine students on cornet-trumpet who auditioned, only twenty-three of each group were recorded as previously planned.

Three recordings of flutes were not of satisfactory quality for the study due to technical problems early in the audition session. Twenty satisfactory recordings were used as the sample for flute auditions, using the three scoring techniques. Recordings of the seventeen students who auditioned on trombones were of good quality, and all were used for the study. The recording of the fourth student on clarinet was not used because of technical difficulties. The twenty-first student recorded on clarinet was used as a replacement for student number four. The first twenty recordings made of cornet-trumpets were of good quality and, therefore, the three additional recordings were not needed.

In summary, the population was composed of a representation of band students from a large geographic area of Texas. Both large and small high schools were represented. The fact that fewer students than were expected auditioned on flute and trombone automatically limited the sample for these two instruments. The sample obtained, however, was recorded during live auditions for all-region band.

V. RECORDING AND EDITING TECHNIQUES

Recording techniques. The high quality of tape recordings which were made at the all-region auditions was considered to be the main factor necessary for proper evaluation of audition performances. As stated previously, the pilot study provided a testing ground for four types of tape recorders. The author felt that only tape recorders and microphones with a frequency response from 50-15,000 Hz at 7 1/2 ips or better should be used for the study. Immediately prior to their use, tape recorders were adjusted to meet original specifications by a franchised Ampex repair center.

Three tape recorders were used for the study. These three were: (a) Ampex stereo tape recorder, AG-500-2, (b) Ampex tape recorder 600, and (c) Sony stereo tape recorder 530. Two types of microphones were used with the tape recorders. These types were: (a) Neumann condenser microphones, Model U-64-2 and (b) Electro-Voice cardioid microphones, Model 666. Since these microphones are used by the East Texas State University Music Department for stereo recording, two microphones of each type were available for the study. The type of recording tape used with these recorders was Century 828, which is the type used in making the tape master for custom records. This tape corresponds in quality with Scotch 201 series tape.

One purpose of the recording procedure was to record auditions without students being aware of it. This was accomplished by hiding the microphones in flower arrangements which were placed in front of the room where students were being recorded. The microphone cable was placed under the inside edge of the door facing and taped to the bottom edge of hallway lockers with gray "duct" tape. The microphone cable was connected to the tape recorders which were located in the adjacent room.

Microphones, microphone cables, and tape recorders were in place by 7:00 a.m., in order for the equipment to be in place before the students arrived about 7:30 a.m. Doors to the rooms containing the recorders were locked from the outside, and the recording technicians did not leave the rooms during auditions. The recording level was estimated before the first student played. Each student was asked by the adjudicator to play a few warm-up notes before he played the audition material, and slight adjustments were made during warm-up notes of each participant.

The recording technician was not able to visually observe the student who was auditioning. The recorder was started as the student entered the room and was stopped as the student left the room. These two actions were not difficult to monitor aurally. The recording technician recorded those students chosen by the random selection process. Due to a smaller number of students than expected, all the flutes and trombones were recorded. A check list was used by the recording technician in order to keep a careful record of students who auditioned. Every student was accounted for regardless of whether he was recorded on tape or not. Tape storage boxes were carefully labeled, and the tapes were returned to the box after the maximum number of students had been recorded on the tape. If the recording technician was doubtful about the length of tape remaining on a reel, he used a new reel of tape for the next student. This precaution was taken to insure that all of the audition would be recorded.

The above plan for the recording arrangement permitted recordings to be made without students being aware of the presence of microphones or recorders. The arrangement of the rooms for auditions provided a maximum amount of efficiency in arranging microphones and recording equipment, and most of the band directors indicated they had forgotten the auditions were being recorded.

Editing techniques. The tapes which were recorded at the all-region festival contained all sounds in the auditioning room from the time the student entered the room until completion of the audition. Only the actual performance of the audition was needed for analysis. In review, the pilot study revealed that the best method for editing tape was: (a) to cut out all tape which did not contain the performance of audition music, (b) to splice a two-foot strip of blank tape between each prepared exercise and sight reading exercise, and (c) to splice a three-foot strip of paper leader tape between the recording of each student. (White paper leader tape was selected in order to label each student's tape for identification.) Each student's tape was labeled with the original recording order number.

The tapes were placed in a different random sequence each time they were ranked. The random sequence for the four ranking steps is shown in Table I. (See page 48.) The order

TABLE I

RANDOM SEQUENCE USED FOR THE ORDER OF TWENTY STUDENTS RECORDED ON EACH OF FOUR INSTRUMENTS (FLUTE, CLARINET, CORNET-TRUMPET, AND TROMBONE) AT ALL-REGION BAND AUDITIONS FOR REGION II AND REGION IV OF STATE OF TEXAS, DECEMBER, 1968

Actual Ranking	Audition Session I	Audition Session II	Audition Session III
7	15	13	11
9	5	6	1
16	4	2	16
15	6	19	17
11	20	16	13
17	11	8	3
20	17	12	20
14	3	14	7
19	14	7	14
12	1	4	12
10	16	17	5
3	7	9	10
18	8	1	15
4	13	5	8
13	9	20	18
8	12	10	6
6	18	11	2
1	10	15	19
5	2	3	9
2	19	18	4

of the tapes was changed by winding each student's tape on an empty five-inch reel, and each reel was assigned a number to correspond with the number of each student's tape. Tapes of the students were then spliced into place on a seven-inch reel in the random sequence prescribed for the next step. Although the order of random sequence differed for each of the four ranking steps listed in Table I, the same random sequence was used for each of the four instruments tested.

VI. DATA COLLECTION

Preliminary steps to data collection. Before the data collection process began, some preliminary steps were necessary. The author traveled to the site of the Region IV (North zone) all-region band auditions to examine the rooms to be used for auditioning. The all-region band organizer was high school band director at the school where auditions were to be held. He made room assignments for auditions and attempted to assign rooms which would be most convenient for recording purposes. Only flute, clarinet, and cornet-trumpet auditions were to be recorded at Region IV. The arrangement of classrooms that were being used for auditions allowed two tape recorders to be placed in a room between the rooms assigned for flute and clarinet auditions. The room for cornet-trumpet auditions was assigned to permit the tape recorder to be located in an adjacent room.

Trombone auditions were recorded in Region II at a later date. The all-region band organizer for that region assigned trombone auditions to be held in the high school library. The tape recorder was located in a large workroom which was behind the library check-out counter. This arrangement permitted the tape recorder to be as near auditions as possible and facilitated concealment of the microphone in the audition room as well as camouflaging of the microphone cable.

The next preliminary step was to test several playback systems. The best portable playback system tested consisted of an Ampex 600 professional recording and playback unit, a Fisher X-101-B amplifier and a high quality eight-inch speaker in a bass-reflex enclosure. This equipment was used for playback of audition tapes in the twelve adjudicating sessions. The performance of each professional musician, which was the variable to the third scoring technique, was played back on a Sony 355 tape deck and an Ampex AA620 speaker. All amplifier settings were pre-determined and were pre-set before each audition session.

The pilot study revealed that a constant factor needed to be used in evaluating differences in the consistency level of the scoring sheets. After recordings were made and edited, the author took the tapes to four high school band directors who had many years of adjudicating experience. These men were asked to evaluate performances on the tapes and to rank the students from best player to poorest player. The rank order which they submitted was to be, to the best of their judgment, the actual rank order of students who were recorded at the all-region band auditions. Each director kept the tapes approximately one week. This rank order was used as a constant factor in the Spearman rho analysis. Although the actual rank order of these twenty students could never be said to be 100 per cent accurate, for the purposes of this research the ranking process described above was used as the order which closely approached the actual rank order.

One of the problems described in Chapter I was the difficulty adjudicators had in establishing a standard for the group which they were evaluating. In order to establish a standard for comparison with the audition tapes, the author asked four professional instrumentalists on each of the four instruments used in the research to record the prepared exercises which were required for the all-region audition. The same recorders and microphones were used to record the professionals as were used for recording students. The recording of the professional instrumentalists was used with the third scoring session for each instrument.

The following chart indicates the location, date, and time for the three audition sessions for flute, clarinet, cornet-trumpet, and trombone:

Instrument	Location	Date	Time
	Audition Session I		
Flute	East Texas State University	12-20-68	4:00- 6:00 p.m.
Clarinet	University of Texas at Arlington	12-21-68	9:00-11:00 a.m.
Cornet-trumpet	Texas Christian University	12-18-68	4:00- 6:00 p.m.
Trombone	Baylor University	12-19-68	3:00- 5:00 p.m.
	Audition Session II		
Flute	East Texas State University	1- 7-69	4:30- 6:30 p.m.
Clarinet	University of Texas at Arlington	1-17-69	4:00- 6:00 p.m.
Cornet-trumpet	Texas Christian University	1-15-69	4:00- 6:00 p.m.
Trombone	Baylor University	1-16-69	3:00- 5:00 p.m.
	Audition Session III		
Flute	East Texas State University	1-21-69	4:30- 6:30 p.m.
Clarinet	University of Texas at Arlington	1-31-69	4:00- 6:00 p.m.
Cornet-trumpet	Texas Christian University	1-29-69	4:00- 6:00 p.m.
Trombone	Baylor University	1-30-69	3:00- 5:00 p.m.

Data collection. The data collection procedure consisted of three audition sessions on each of four band instruments: flute, clarinet, cornet-trumpet, and trombone. The reason for three sessions was to give the author an opportunity to test three techniques of evaluation. The scoring sheets consisted of a continuum for each of six performance characteristics. The first scoring sheet contained numbers which were printed above the continuum. The scoring sheet for the second session contained verbal descriptions of the performance characteristics above the continuum. (See Appendix B.) In the third audition session, a standard of performance for the audition was established by playing the tape recording of the professional instrumentalist for adjudicators before each student was heard. Student performances were scored using a scoring sheet which was identical to the one used in the second audition session.

Playback equipment was positioned at the front of the room where audition sessions were held. Adjudicators were seated at tables approximately fifteen feet from the playback equipment. Equipment was set up approximately forty-five minutes before each session was to begin. The playback mode of the recorder was tested, and amplifier controls were set in the pre-determined position. The recorder and amplifier were left in "on" position to allow the electronics to warm up properly before the audition was to begin.

Scoring sheets prescribed for each audition session were placed in three file folders labeled Judge 1, Judge 2, and Judge 3, and each of the folders contained the exact number of scoring sheets needed. The scoring sheets in each folder were numbered consecutively, and one scoring sheet was used to score each student. Copies of the music for prepared exercises and sight reading were placed on music stands positioned adjacent to tables where the judges were seated. Instructions to judges were recorded on tape by the author and played before each audition session began. (See Appendix D for the text of these instructions.) When all students had been heard and evaluated, the judges were asked to place the scoring sheets back in the folders. Each set of scoring sheets was identified by placing the name of the instrument and the number of the judge in the upper left-hand corner of the scoring sheet for the first student. The set of scoring sheets was clipped together and placed in the folders for later scoring.

One question which was asked by adjudicators in the first audition session was: "What standards should we use in evaluating the recordings?" The author responded that adjudicators were to use their own personal standards in the evaluation of each student. Another question which was asked was: "Are we to place the marks under numbers or at any point along the continuum?" The author responded that adjudicators were to place marks at any point on the continuum. In order to expand the understanding of why the mark could be placed at any point, the explanation was given that a transparency with a value scale would be placed over scoring sheets, and a proportionate value would be given to a mark at any point on the continuum. In the second audition session adjudicators again asked the question concerning where the mark was to be placed on the continuum, and again they were informed that the mark could be placed at any point on the continuum. Since the technique of scoring the sheets was the same for the third audition session, no questions were asked by adjudicators.

Transparencies of a value scale were processed and used to score adjudication sheets from the twelve audition sessions. (See Appendix E for the value scale.) A transparency of the value scale was placed over each scoring sheet, and the value of the mark across the continuum was determined for each of the six performance characteristics. The six scores were summed, and the rank order was determined, i.e., the highest score was number one, the second highest was number two, etc. The data in the form of the rank order of the audition was punched on IBM cards, and the statistical computations performed on an IBM Model 360 computer.

VII. SUMMARY

A survey of present methods of evaluation in festival band auditions revealed that a percentage range was the major guideline for the adjudicator to use in his evaluation. The scoring sheet used in the Band Camp audition was more systematic, but it also required the adjudicator to transfer a subjective picture of his evaluation to a specific number. The application of the continuum to the scoring of a musical performance seemed to have considerable merit. The lack of a related research project created a necessity for experimentation and testing of all procedural steps. Recording and editing techniques required careful evaluation to insure high quality recordings for the auditions.

CHAPTER IV

ANALYSIS OF RESULTS

An analysis of data obtained in this research is the purpose of this chapter. Presentation of the data is divided into four sections: statistical methods, interjudge reliability, rank-difference correlation, and interjudge reliability of performance characteristics. Adjudicators are not identified by name but are referred to as A, B, and C. The "A" designates college band directors, "B" designates applied music teachers or professional performers, and "C" denotes high school band directors. The subscripts, "w" and "e," are used to identify data obtained with the weighted value scale and the equal value scale, respectively.

I. STATISTICAL METHODS

The statistical analysis of data in this report was intended to describe the degree of reliability and validity obtained by the three audition techniques. In practice, the important decision which the audition must reflect was the rank order of students who auditioned. For this analysis, the Spearman Rank-Difference Correlation rho was applied to data. Mueller and Schuessler gave the rationale for using this formula by presenting the following analysis:

> . . . it is evident that rho measures the correlation between ordinal ranks, and not the correlation between

> potential magnitudes that are being ranked. Hence, rho in general overstates the degree of congruity existing between the raw variates, expressed or unexpressed. Thus, two judges ranking the same exhibit of pictures may be at wholly different locations on the complete continuum of taste preference, but still rank the works of art in identical order. . . . it is reasonable to suppose that similar rankings do correspond to similar preferences, owing to the fact that by and large individuals share a common culture.[1]

The formula for computing the Spearman Rank-Difference Correlation rho is:

$$P = 1 - \frac{6 \Sigma D^2}{N (N^2 - 1)}$$

Where D = difference between paired ranks

N = number of students ranked

Guilford expressed confidence in this statistical formula when correlations are high. He wrote: ". . . we may have almost as much confidence in rho for indicating the amount of correlation as we have in [Pearson's] r applied to samples of the same size."[2]

The interjudge reliability of ratings was determined by Kendall's Coefficient of Concordance W. This procedure is described by Downie and Heath as a statistical formula which can be applied when measurement data are reduced to ranks, when the samples are small, and when assumptions for parametric statistics cannot be met.[3] The relationship of the ranks for the three adjudicators was determined with the following formula:

---

[1] John H. Mueller and Karl F. Schuessler, Statistical Reasoning in Sociology (Boston: Houghton Mifflin Company, 1961), pp. 274-275.

[2] J. P. Guilford, Fundamental Statistics in Psychology and Education (New York: McGraw-Hill Book Company, 1965), p. 308.

$$W = \frac{12 \sum D^2}{m^2 (N) (N^2 - 1)}$$

Where m = number of ranks

D = difference between ranks

N = number of students ranked

This statistical formula was applied to the ranks of the summed scores for each individual and also used to compute the degree of interjudge reliability in the ranks derived from the six performance characteristics.

II. COMPARISON OF AUDITION DATA

Interjudge reliability. The analysis of the interjudge reliability of the three auditioning techniques revealed that the reliability coefficient of all scores was significant at the .01 level of confidence. This was accomplished by transforming Kendall's Coefficient of Concordance W into an F score. The formula for this transformation was as follows:

---

[3] N. M. Downie and R. W. Heath, Basic Statistical Methods (New York: Harper and Row, Publishers, 1965), p. 206.

$$F = \frac{(m - 1)\,W}{1 - W}$$

The F score was interpreted by using an F table which appears in most basic statistics books, i.e., Guilford, Table F.[4] An analysis of W in Table II (See page 61.) indicated that the range of scores in Audition Session II was smaller than the range in Audition Session I or III. In the case of cornet-trumpet scores and trombone scores, Audition Session II was higher than Audition Session I or III. The interjudge reliability of flutes remained almost constant during all audition sessions. Clarinets improved the reliability coefficient a small degree with each session.

Rank-difference correlation. The purpose of this statistical procedure was to determine the correlation of each audition technique when compared with an actual ranking by an expert adjudicator. By comparing the rank order of each audition technique with the constant factor of the actual ranking, the coefficient of correlation between the two rankings indicated the degree of consistency of the three audition techniques with the actual ranking. The rationale of this method was that a large increase of correlation coefficient would indicate that adjudicators who heard the tape recordings of the auditions only one time for each audition session were

---

[4] Guilford, op. cit., pp. 583-586.

TABLE II

KENDALL'S COEFFICIENT OF CONCORDANCE W FOR INTERJUDGE RELIABILITY OF THREE ADJUDICATORS

Audition Session	Flute	Clarinet	Cornet-Trumpet	Trombone
I_w	.89	.82	.79	.79
I_e	.89	.82	.79	.78
II_w	.89	.84	.86	.92
II_e	.88	.84	.83	.91
III_w	.91	.91	.77	.82
III_e	.89	.92	.77	.82

NOTE: The subscript "w" signifies the score by the weighted value scale. The subscript "e" signifies the score by the equal value scale.

able to more closely approach a ranking of the same tapes which was determined after a week of evaluation by an expert adjudicator.

The coefficients of correlation were all significant at the .05 level of confidence, and all but one of the weighted value scores and one of the equal value scores were significant at the .01 level of confidence. Tables III-VI provide an analysis of the data computed from rho for the three audition sessions of the four instruments. (See pages 63-66.)

One additional audition session was held to determine if a different adjudication panel would confirm the reliability coefficient and the correlation coefficient of the previous audition sessions. For this audition session, the second scoring technique (verbal concept) was used because in previous audition sessions the range of the W coefficients was not as wide as the other two scoring techniques, and the rho coefficients were all significant at the .01 level of confidence. Also, the adjudicators seemed to prefer the verbal description concept.

The tapes for flutes were used for two reasons. First, the interjudge reliability of the three adjudicators who evaluated flutes remained at the most consistent level of the four adjudication panels used in the research. Secondly, the author was able to obtain the use of the room where the previous audition sessions for flutes had been conducted; therefore,

TABLE III

SPEARMAN'S RANK-DIFFERENCE CORRELATION *rho* OF ADJUDICATORS AND ACTUAL RANKINGS FOR *FLUTES*

Adjudicator	Audition Session I	Audition Session II	Audition Session III
A_w	.80	.72	.81
A_e	.81	.72	.83
B_w	.75	.76	.74
B_e	.75	.75	.78
C_w	.71	.82	.69
C_e	.71	.80	.70

NOTE: The subscript "w" signifies the score by the weighted value scale. The subscript "e" signifies the score by the equal value scale.

TABLE IV

SPEARMAN'S RANK-DIFFERENCE CORRELATION rho OF ADJUDICATORS AND ACTUAL RANKINGS FOR CLARINETS

Adjudicator	Audition Session I	Audition Session II	Audition Session III
A_w	.64	.83	.76
A_e	.69	.83	.77
B_w	.70	.87	.85
B_e	.67	.86	.83
C_w	.66	.79	.83
C_e	.64	.80	.80

NOTE: The subscript "w" signifies the score by the weighted value scale. The subscript "e" signifies the score by the equal value scale.

TABLE V

SPEARMAN'S RANK-DIFFERENCE CORRELATION *rho* OF ADJUDICATORS AND ACTUAL RANKINGS FOR CORNET-TRUMPETS

Adjudicator	Audition Session I	Audition Session II	Audition Session III
A_w	.68	.75	.40
A_e	.66	.74	.38
B_w	.81	.62	.60
B_e	.81	.66	.64
C_w	.65	.53	.69
C_e	.68	.53	.72

NOTE: The subscript "w" signifies the score by the weighted value scale. The subscript "e" signifies the score by the equal value scale.

TABLE VI

SPEARMAN'S RANK-DIFFERENCE CORRELATION *rho* OF ADJUDICATORS AND ACTUAL RANKINGS FOR TROMBONES

Adjudicator	Audition Session I	Audition Session II	Audition Session III
A_w	.77	.94	.83
A_e	.79	.93	.83
B_w	.63	.93	.82
B_e	.68	.91	.81
C_w	.73	.90	.88
C_e	.72	.91	.88

NOTE: The subscript "w" signifies the score by the weighted value scale. The subscript "e" signifies the score by the equal value scale.

variables were held to a minimum. The same tapes in the order of the second session were evaluated in the same room, on the same playback equipment, at approximately the same time of day. The only change was the use of a different adjudication panel.

The results of this additional audition session revealed that the interjudge reliability or W was $.87_w$ and $.85_e$ using the weighted and equal value scale respectively. The correlation coefficient with the actual ranking was as follows:

	Adjudicator		
	A	B	C
Weighted value scale	.79	.76	.68
Equal value scale	.77	.76	.69

Interjudge reliability of performance characteristics. An analysis of the interjudge reliability was made in order to determine if any single characteristic was more reliable than the others. The score of each student for each characteristic, i.e., tone, pitch, rhythm, etc., was ranked. The rankings of the three judges on each characteristic were computed using Kendall's Coefficient of Concordance W. The coefficient obtained by W was transformed into an F score in order to determine the degree of significance. All the scores for the characteristics of technical accuracy and musicality were significant at the .05 level of confidence. All but two scores in musicality were significant at the .01 level of confidence. Figures 1-6 picture graphically the degree of variation present in an analysis of the six performance characteristics. (See pages 68-73.)

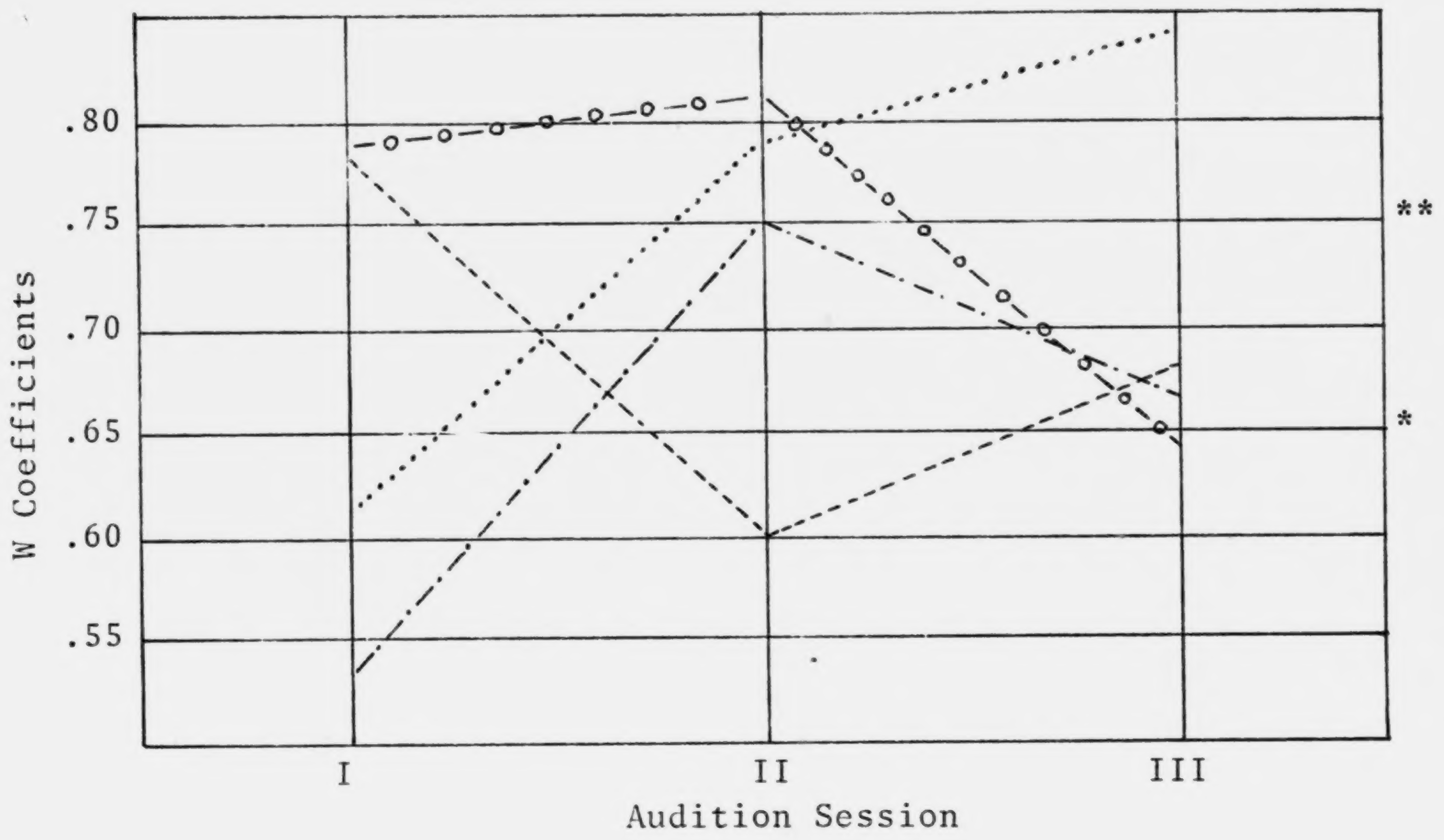


FIGURE 1

A COMPARISON OF THE W COEFFICIENTS FOR TONE

** Scores above this point are significant at the .01 level of confidence
* Scores above this point are significant at the .05 level of confidence

FLUTE ....
CLARINET ----------

CORNET-TRUMPET -·-·-·-·-·-·-
TROMBONE -o-o-o-o-o-o-o-

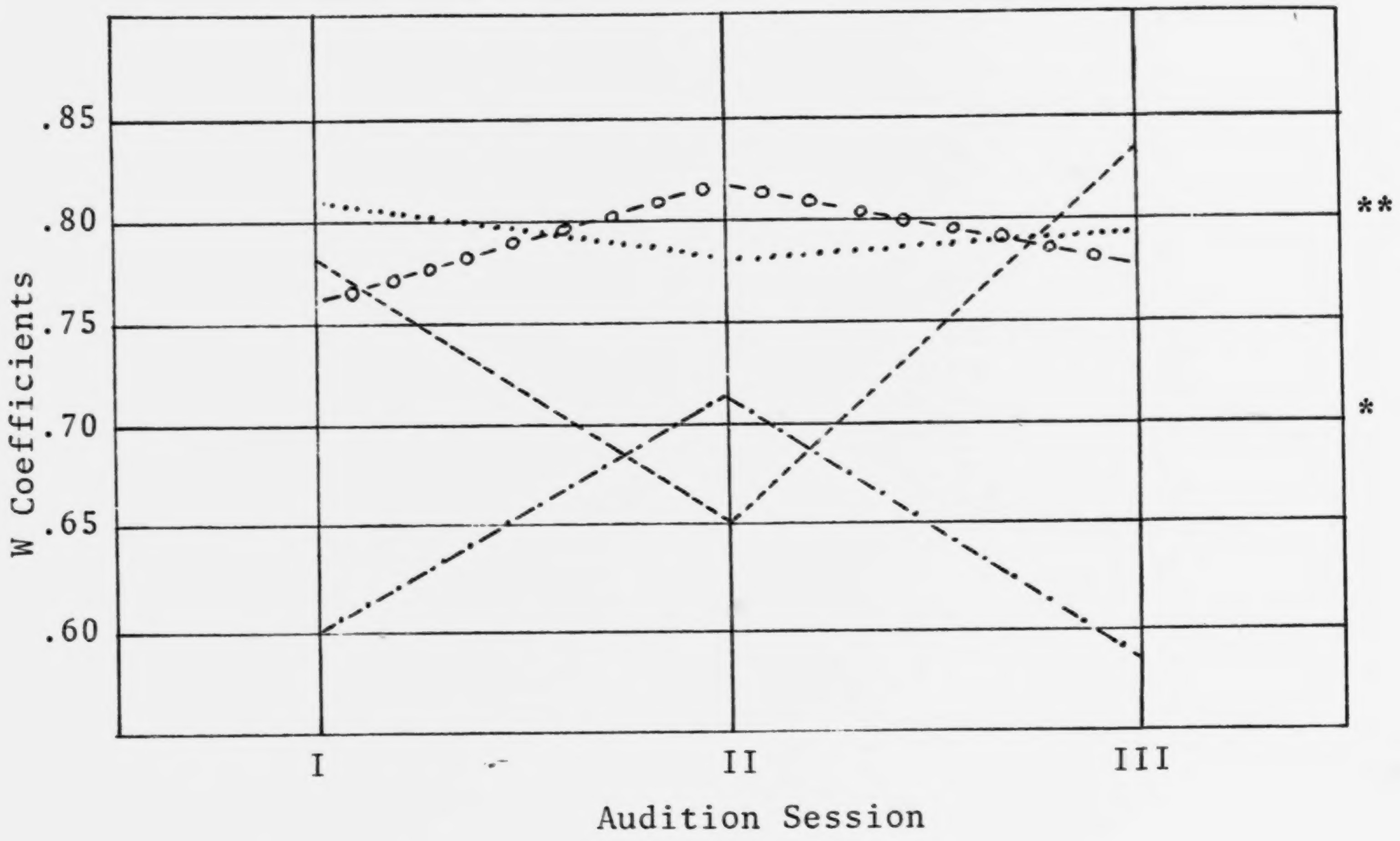


FIGURE 2

A COMPARISON OF THE W COEFFICIENTS FOR PITCH

** Scores above this point are significant at the .01 level of confidence
* Scores above this point are significant at the .05 level of confidence

FLUTE ....
CLARINET ----------

CORNET-TRUMPET -·-·-·-·-·-·-
TROMBONE-o-o-o-o-o-o-

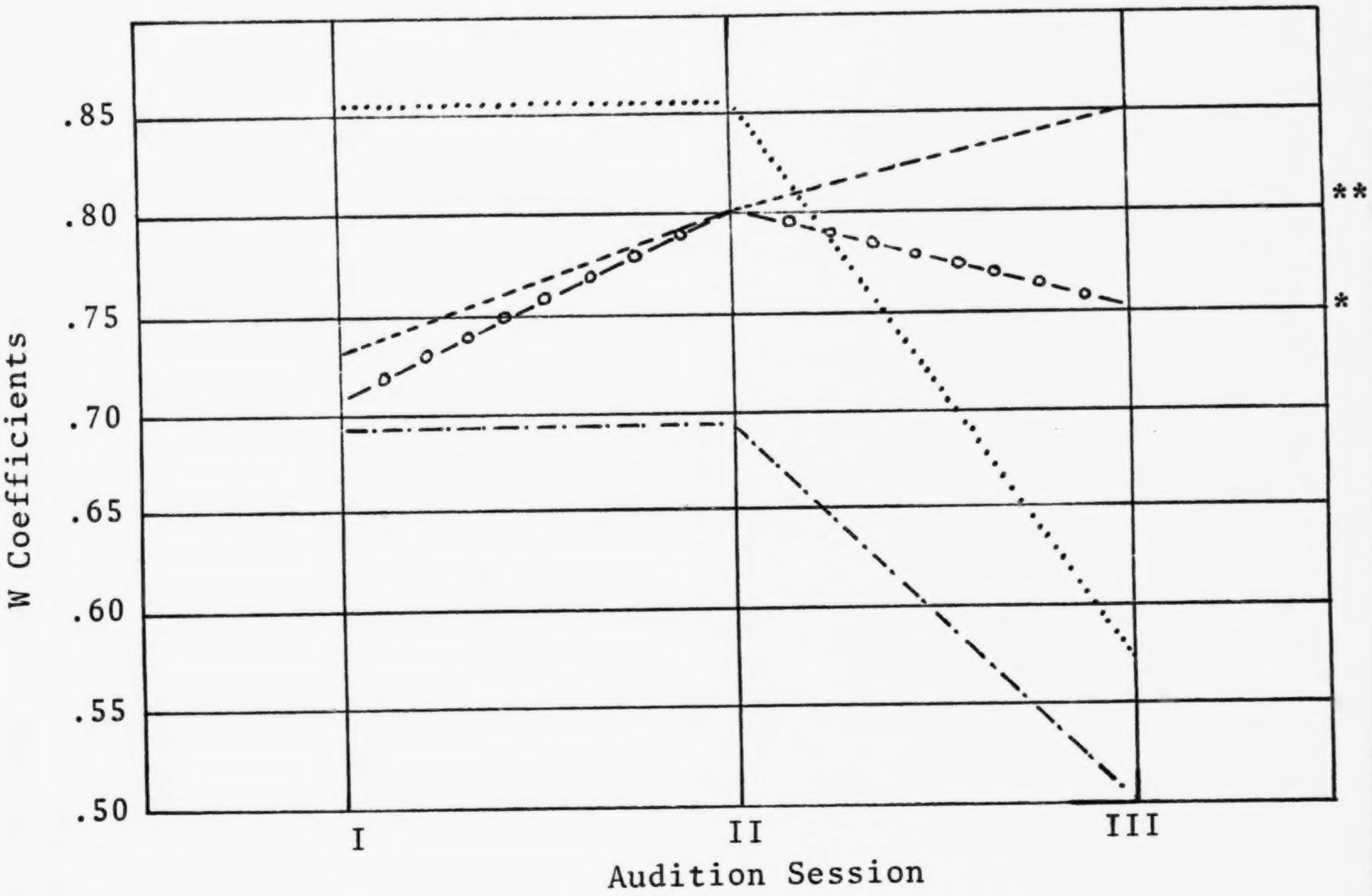


FIGURE 3

A COMPARISON OF THE W COEFFICIENTS FOR RHYTHM

**Scores above this point are significant at the .01 level of confidence
* Scores above this point are significant at the .05 level of confidence

FLUTE ....
CLARINET ----------

CORNET-TRUMPET -·-·-·-·-·-·-
TROMBONE -o-o-o-o-o-o-

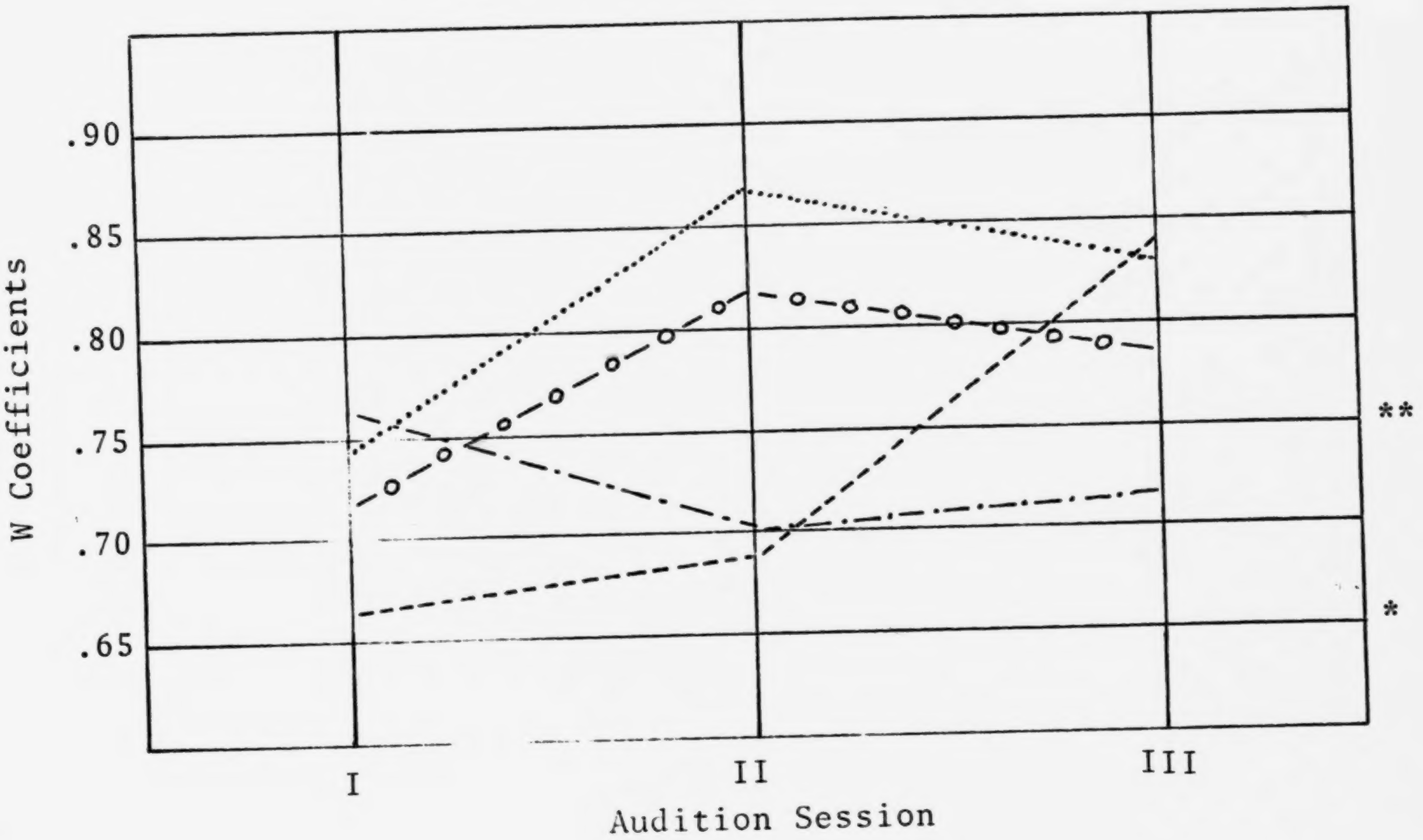


FIGURE 4

A COMPARISON OF THE W COEFFICIENTS FOR TECHNICAL ACCURACY

**Scores above this point are significant at the .01 level of confidence
* Scores above this point are significant at the .05 level of confidence

FLUTE ....
CLARINET ----------

CORNET-TRUMPET -·-·-·-·-·-·-
TROMBONE -o-o-o-o-o-o-

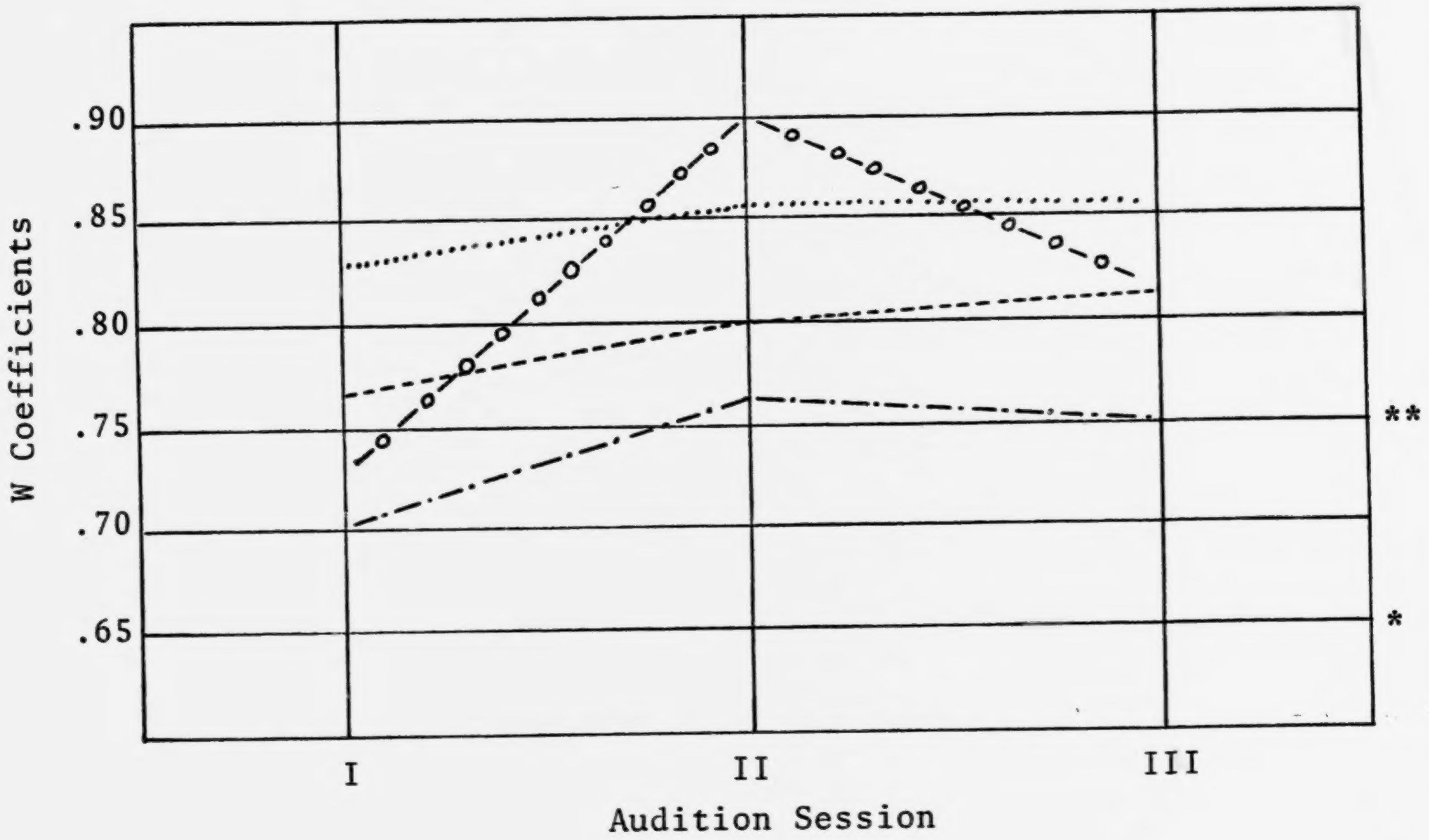


FIGURE 5

A COMPARISON OF THE W COEFFICIENTS FOR MUSICALITY

**Scores above this point are significant at the .01 level of confidence
* Scores above this point are significant at the .05 level of confidence

FLUTE ....
CLARINET ----------
CORNET-TRUMPET -·-·-·-·-·-·-
TROMBONE -o-o-o-o-o-o-

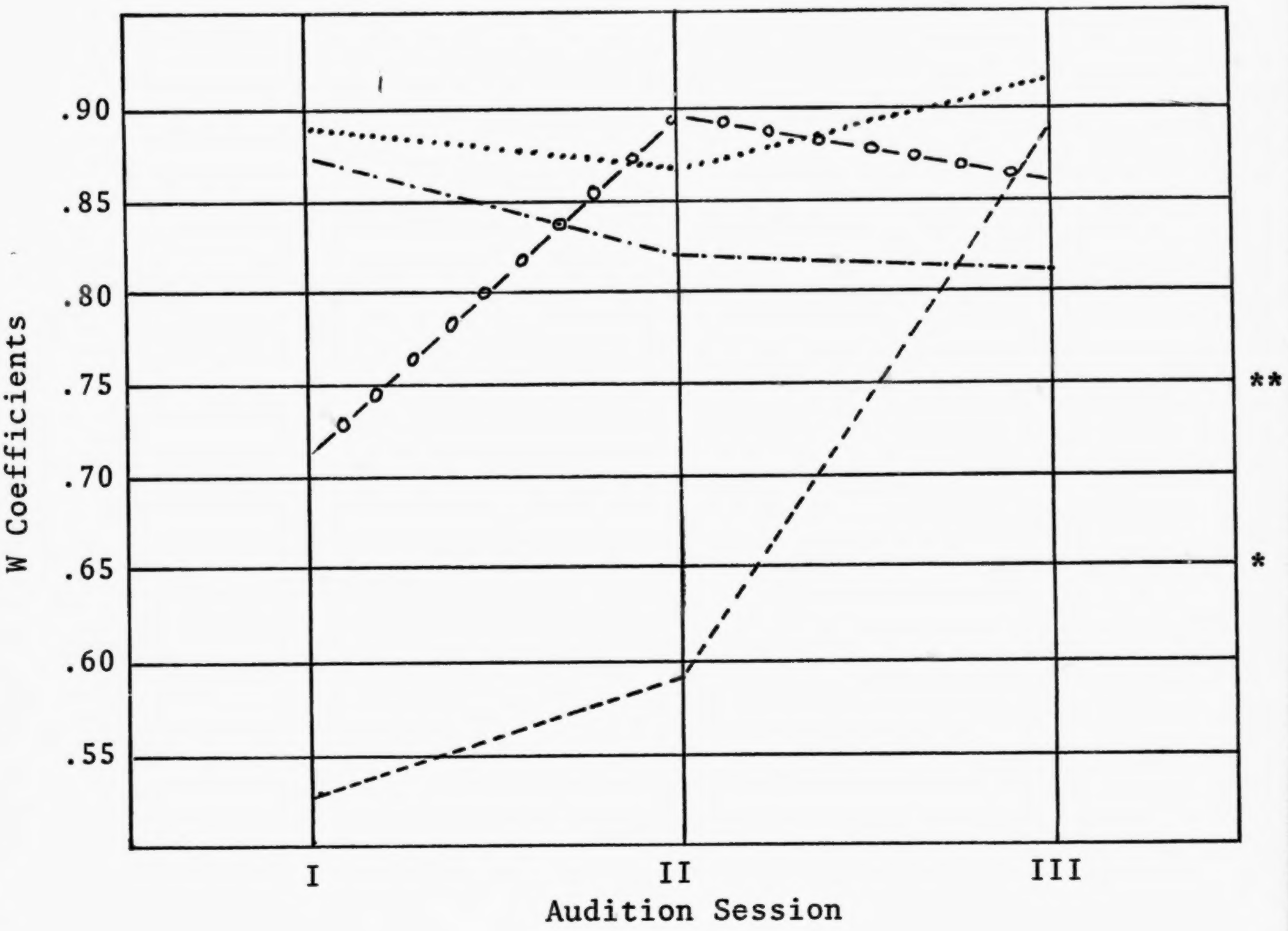


FIGURE 6

A COMPARISON OF THE W COEFFICIENTS FOR SIGHT READING

**Scores above this point are significant at the .01 level of confidence
* Scores above this point are significant at the .05 level of confidence

FLUTE ....
CLARINET ----------
CORNET-TRUMPET -.-.-.-.-.-.-
TROMBONE -o-o-o-o-o-o-

The additional audition session for flutes with a different panel of adjudicators revealed that the interjudge reliability or W for technical accuracy was .85 and the W coefficient for musicality was .83. Both scores are significant at the .01 level of confidence.

CHAPTER V

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

I. SUMMARY

The need for a more reliable method of musical performance evaluation was the motivating factor of this research from the early stages of planning to the final stages of data collection. Since research projects directly related to the subjective evaluation of musical performance had not been conducted and the results disseminated, it was necessary that procedural techniques be developed and tested under live conditions in a pilot study in order to properly evaluate their effectiveness. Although this research was only an initial step in the study of criteria for evaluation of musical performance auditions, audition techniques developed by this study should give persons who administer auditions an indication of the reliability and validity of these techniques.

General summary. This research was designed to provide music educators who are involved in musical performance auditions an evaluation of audition techniques which seem to offer a more reliable evaluation of musical performance. The pilot study, which was conducted during the Blue and Gold Band Camp on the campus of East Texas State University in June 1968, provided the opportunity to test several different tape recorders and several recording and editing techniques. A survey of methods currently being used in Texas all-region band auditions indicated that a percentage range was ascribed to different performance areas and that the adjudicator transferred his subjective evaluation to a number within the allowed percentage range.

Tape recordings of students who were recorded at the Band Camp were evaluated, using three audition techniques. Two scoring sheets were developed and used for the first two audition sessions which were based on numerical and verbal description concepts which have been used for rating in many areas, i.e., musical performance, merit rating by management, attitude scales, and other subjective evaluations. In the third audition session a recording of a professional instrumentalist was used in order to provide a standard for adjudicators. This concluded the pilot study portion of the research.

Audition techniques were refined as a result of the pilot study. Students who auditioned for all-region band in Region II and the North zone of Region IV (divisions designated by Texas Music Educators Association) on flute, clarinet, cornet-trumpet, and trombone were recorded, and tapes were edited for use in three audition sessions. Adjudicators were selected, and dates for audition sessions were confirmed. The tape recordings were delivered by the author to the adjudicators who were to give the actual ranking of students who were recorded. The tape recordings were then used in the three audition sessions, and the rank order was determined when scoring sheets marked by each judge were evaluated. Statistical analysis of the rank order provided an interjudge reliability coefficient of the three adjudicators and a correlation coefficient of each adjudicator's ranking with the actual ranking.

Adjudicator reaction. The initial reaction of the adjudicators was a feeling of surprise that they would not know the rank order of students after the tape recordings had been evaluated. The fact that they were to only evaluate the musical performances on the tape recordings seemed to bother them at first. They were anxious to learn which performance they had evaluated as the best. Some adjudicators seemed to hesitate to use the entire continuum in their evaluation. The fact that they were not familiar with the use of a continuum in musical performance evaluation caused them to place the evaluation mark directly under a number. For example, one adjudicator rated the tone of nine students the same in the first audition session. Another adjudicator expressed his preference of the scoring sheet (numerical concept) which was used in the first audition session.

The scoring sheet used in the second audition session (verbal concept) was preferred by the majority of the adjudicators. They seemed to feel that the verbal descriptions helped them to picture different degrees of musical performance along the continuum better than did numbers. The one adjudicator who preferred numbers as a guide indicated that verbal descriptions seemed to restrict his personal standard. The scores of adjudicators who preferred the second technique were higher when they used this technique. Reaction to the wording of the verbal descriptions was favorable, and no changes were suggested.

In the third audition session, the adjudicators seemed to be confused by the tape recording of the professional instrumentalist. The profile of the scores indicated that the degree of this interference with their own standards caused the reliability coefficient to be lower in cornet-trumpet and trombone auditions. Although the clarinet adjudicators indicated a degree of confusion, their reliability coefficient was higher for Audition Session III.

A majority of the adjudicators preferred the scoring sheet used in the second audition session. However, all were of the opinion that use of the continuum could be effectively applied to evaluation of musical performance. This positive reaction of the adjudicators after they had used the continuum concept as applied to musical performance evaluation was in contrast with their initial feelings of mild skepticism.

II. CONCLUSIONS

General conclusions. To the extent that the facts obtained and the opinions expressed are accurate and insofar as the population included in the study are representative of the whole, the following conclusions may be drawn as of the time this study was made:

1. Since the statistical analysis revealed that all the data were significant at the .05 level of confidence, any one of the three audition techniques can be used with an equal degree of confidence. However, the adjudicators seemed to prefer the verbal descriptive continuum. Although not statistically more significant than the other two, the pattern of scores for the second audition technique seemed to indicate that it might be more consistent in its ranking than either audition technique one or three.

2. The evaluations of adjudicators were scored effectively without the need for their evaluations to be transferred to a numerical score by the adjudicators themselves. The adjudicators were able to become evaluators of musical performance in a reliable manner without immediately knowing where they had placed students in the rank order. The concept of the adjudicator as an evaluator only, may have a secondary effect. It may relieve some pressure on the adjudicator in auditions for festival bands.

3. The scoring sheets provided a simple and quick method for marking the evaluations for ranking purposes. Adjudicators were able to mark their evaluations of the six performance characteristics without the difficulty of a numerical transformation.

4. The actual scoring and tabulating of results to determine rank order of students who were evaluated on the graphic scoring sheets were time-consuming. Although the adjudicators were able to mark scoring sheets quickly, the process of scoring the evaluation proved to require more time than the methods presently being used.

5. Musicality was the most significant performance characteristic when analyzed for interjudge reliability. The characteristic of musicality was not as significant as the composite score of the six characteristics tested, but it more closely approached significance than did the other characteristics.

6. The adjudicators seemed less fatigued after evaluating auditions where three or four medium-length exercises, i.e., twenty-four measures, were used than in audition sessions where one or two long exercises, i.e., forty-eight measures, were used. The two exercises for cornet-trumpet auditions were long, and adjudicators seemed to be more restless during the second half of the audition session, especially if the student's performance was poorly prepared. Trombone and clarinet exercises were shorter and seemed to hold the interest of adjudicators through the duration of the audition session.

7. The degree of difference in performance between students who auditioned seemed to be in direct relationship with the interjudge reliability of the three judges. For

example, many students who auditioned on cornet-trumpet and clarinet seemed to be almost identical in performance ability. Performances on flute and trombone seemed to be more definable although the personal preference of the individual adjudicators caused some variation in the rank order.

8. The adjudicator may use any standard of performance as long as he is consistent. The fact that an adjudicator marked the students high or low was not important as long as he maintained the same standard for all students. For example, one adjudicator scored the first student very low. This student was one of the best performers. The adjudicator was able to maintain his consistency within a very limited range of the continuum, and his correlation coefficient with the actual ranking only deviated slightly from the previous one.

9. The additional audition session provided an added indication of the validity of the scoring techniques. By retesting the tape recordings which maintained the most consistent interjudge reliability, with a different adjudication panel, the author was able to show that the high reliability coefficient was the result of factors other than the expertise of the original panel.

III. RECOMMENDATIONS

Recommendations for application. Scoring sheets which were developed in this research can be used in most auditions for festival bands where a rank order must be determined. A possible exception could occur when an all-state position or solo performance is to be determined. In this case, a more careful evaluation might be required by hearing students who ranked high in the initial auditions a second time with a more critical analysis of both music played and level of performance attained.

During the first audition session, one adjudicator thought of an additional refinement which seemed to have merit. He placed a dot above each continuum at the point of his evaluation during the early stages of the performance. As the audition progressed, he would place additional dots above the continuum of each performance characteristic to indicate his evaluation at different times during the audition. When the audition was completed, he would place the mark through the continuum at the point which was an average of several dots he had placed on the sheet earlier. This method seemed to give the student credit for performance at different stages of the audition. This adjudicator's evaluations for the second and third audition sessions were higher. Although an effort was not made to test this idea, the author feels it has merit of a practical nature when evaluating students.

The author is of the opinion that the time required to score, tabulate, and rank students from the scoring sheets would not be a problem if several individuals were assigned to collect scoring sheets and score them at periodic intervals during the audition. Scoring sheets can be coded with a number in the top left-hand corner of the page for the instrument and the student's number in the top right-hand corner in the order of participation. This would increase the amount of administrative organization needed for auditions to a small degree, but the increased reliability of the evaluation by the adjudicator is worth the additional effort. Some festival directors already provide this step in order to eliminate the necessity of having the adjudicator add the scores during auditions.

Before an adjudicator uses either scoring sheet, he should have some instructions concerning the method to be used in marking the continuum and the fact that he is only an evaluator of the student's performance level. The festival chairman should be available to answer any questions of adjudicators which may clarify an erroneous concept. The questions may be of a minor nature, but any misunderstanding of scoring sheets could bias the scoring process. As a result of the pilot study auditions and main auditions of the four instruments, the author feels that only a brief amount of time is required for the adjudicator to adapt himself to the use of these scoring sheets.

The construction of the scoring sheets requires that one sheet be used for each student. In order to reduce the relative nature of subjective evaluation, the author recommends that the adjudicator not compare scoring sheets which have been marked with any one sheet he is marking during the audition. Adjudicators differ concerning the value of rating comparison; however, for this type of scoring such a comparison might bias the evaluation and limit the reliability of the rank order.

Audition techniques described and tested in this study have many applications. The obvious one is in auditions for festival bands. They may also be used by band directors in seating auditions for different sections in their bands. The fact that the adjudicator may use any standard of performance he wishes, as long as he is consistent, permits the audition techniques to be used on any level, from junior high school through college.

Recommendations for further research. The true value of this study can only be realized if the findings are supported or rejected through additional research. One research study in an area where so many questions remain unanswered cannot hope to offer more than a method that seems to provide a better way of evaluating musical performance. Additional research is needed in methods of scoring the subjective evaluation of a musical performance.

Several areas of the auditioning process should be studied. Some of these areas are: (a) What music is best suited for auditions? (b) What are the psychological influences of the atmosphere in which the audition is held?, i.e., Should the adjudicator talk to the student?, Should the student see the adjudicator?, Should the adjudicator see the student? (c) What organizational procedure produces the most reliable results?

SELECTED BIBLIOGRAPHY

SELECTED BIBLIOGRAPHY

A. BOOKS

Downie, N. M. and R. W. Heath. Basic Statistical Methods. New York: Harper and Row, Publishers, 1965.

Lehman, Paul R. Tests and Measurements in Music. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1968.

Guilford, J. P. Fundamental Statistics in Psychology and Education. Fourth edition. New York: McGraw-Hill Book Company, 1965.

__________. Psychometric Methods. Second edition. New York: McGraw-Hill Book Company, 1954.

Mueller, John H. and Karl F. Schuessler. Statistical Reasoning in Sociology. Boston: Houghton Mifflin Company, 1961.

Whybrew, William E. Measurement and Evaluation in Music. Dubuque, Iowa: William C. Brown Company, 1962.

B. PUBLICATIONS OF LEARNED SOCIETIES AND OTHER ORGANIZATIONS

Belfour, C. Stanton. "What Have Contests Done for Music Education?" Music Educators National Conference Yearbook. Chicago: Music Educators National Conference, 1936.

Dykema, Peter W. "Observations of An Adjudicator," Music Supervisors National Conference. Chapel Hill, North Carolina: Music Supervisors National Conference, 1926.

"Historical Sketch--Competition-Festivals," Music Educators National Conference Yearbook. Chicago: Music Educators National Conference, 1938.

Lesinsky, Adam P. "The Festival-Contest," Music Educators National Conference Yearbook. Chicago: Music Educators National Conference, 1934.

McAllister, A. R. "Contest Management," Music Educators National Conference Yearbook. Chicago: Music Educators National Conference, 1936.

Patrick, Nelson G. (ed.) "Minutes of 1938 Annual Meeting," Minutes and Procedures of Texas Music Educators Association. Austin, Texas: Texas Music Educators Association, 1961.

"School Music Competition-Festivals," Music Educators National Conference Yearbook. Chicago: Music Educators National Conference, 1940.

"TMEA, What Is It Doing for Music Education in Texas?" Houston, Texas: Texas Music Educators Association, 1968.

Wilcox, E. W. "The Long View in Music Contests," Music Supervisors National Conference. Chapel Hill, North Carolina: Music Supervisors National Conference, 1927.

C. PERIODICALS

Bachman, Harold B. "Competent Adjudication," The Instrumentalist, XVIII (March, 1963), 50-54.

Neilson, James. "Role of the Adjudicator," The Instrumentalist, X (May, 1958), 22.

Whybrew, William E. "The Adjudicator as Educator," The Instrumentalist, XIII (May, 1959), 53 and 91.

D. UNPUBLISHED MATERIALS

Higgins, Roy L. "An Analysis of Adjudication Procedures for A-AA Bands in University of Texas League Competition," Unpublished Master's thesis, The University of Texas, Austin, 1965.

Patrick, Nelson G. A Handbook for Adjudication. Austin, Texas: The University of Texas Interscholastic League, (undated) (mimeographed).

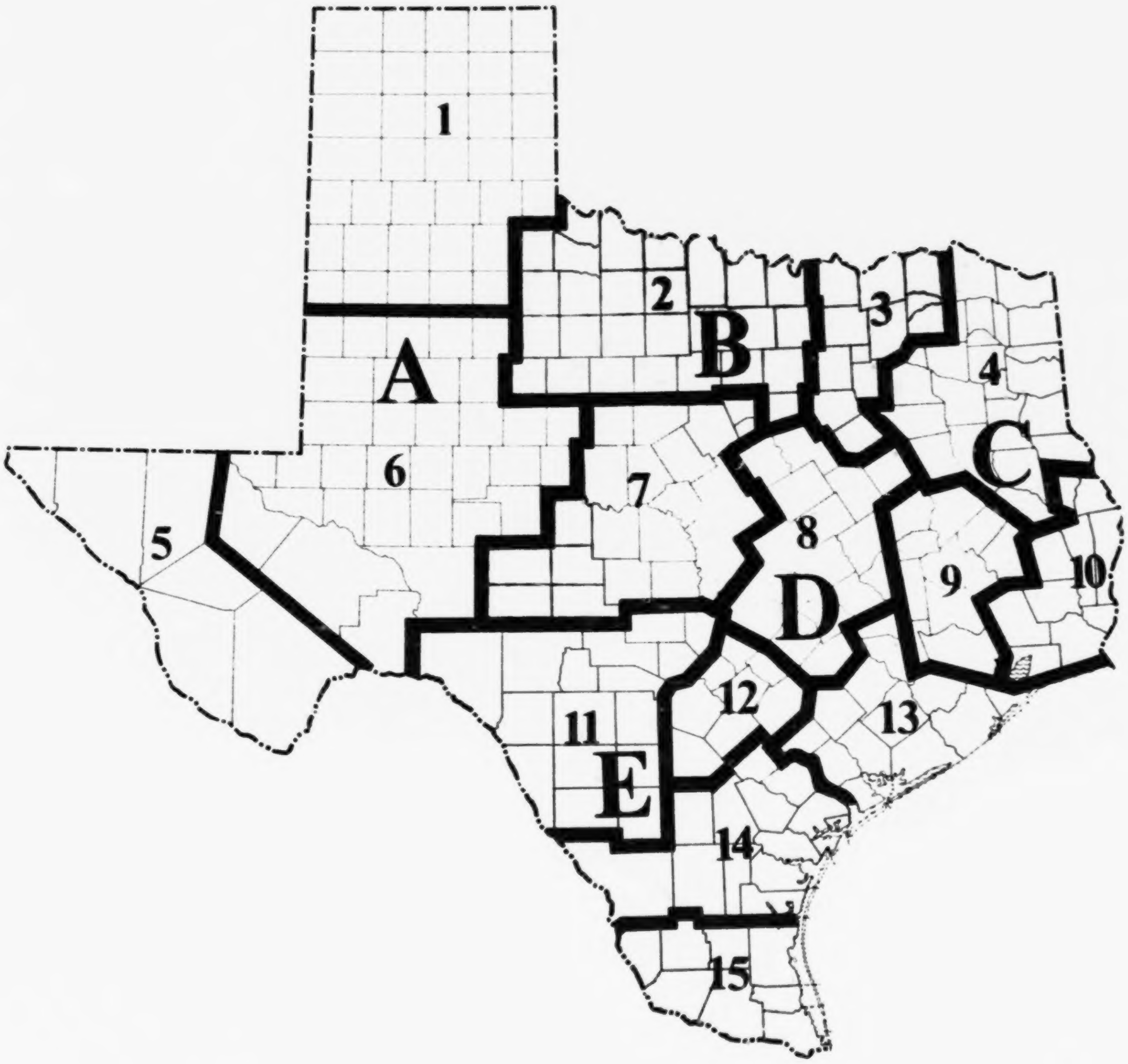
E. OTHER PUBLICATIONS

Watkins, John G. and Stephen E. Farnum. Watkins-Farnum Performance Scale. Winona, Minnesota: Hal Leonard Music, Inc., 1962. For all wind and percussion instruments. Form A and B (equivalent) $6.50 per set (either form), including manual-test book and pad of scoring sheets for various instruments. Administered individually.

APPENDICES

APPENDIX A

A Map of Texas
Indicating Regions and Areas


1
2
B
3
A
4
6
C
7
5
8
9
10
D
12
13
11
E
14
15

A MAP OF TEXAS INDICATING THE
FIFTEEN REGIONS AND FIVE AREAS PRESCRIBED FOR
MUSIC COMPETITION IN THE SCHOOL YEAR 1968-69
BY TEXAS MUSIC EDUCATORS ASSOCIATION

APPENDIX B

Adjudication Scoring Sheets

STUDENT #..........

ADJUDICATION SCORING SHEET

Tone

0 2.5 5 7.5 10

Pitch

0 2.5 5 7.5 10

Rhythm

0 2.5 5 7.5 10

Technical Accuracy

0 2.5 5 7.5 10

Musicality

0 2.5 5 7.5 10

Sight Reading

0 2.5 5 7.5 10

STUDENT #..................

ADJUDICATION SCORING SHEET

Tone

Tone not centered, consistently not characteristic of instrument.	Tone centered on a few notes, many serious tonal problems.	Tone centered some of the time. several problems of tone production.	Tone centered most of the time, a few problems with tone.	Tone centered on every note, quality characteristic of instrument.

Pitch

Consistently out of tune.	Many serious deviations in pitch.	Several serious deviations in pitch.	A few deviations in pitch.	Very few or no deviations in pitch.

Rhythm

No feeling of rhythmic note value.	Many serious deviations in rhythm.	Several serious deviations in rhythm.	A few deviations in rhythm.	Very few or no deviations in rhythm.

Technical Accuracy

Inaccurate in every technical effort.	Many serious technical flaws.	Several serious technical flaws.	A few technical flaws.	Very few or no technical flaws.

Musicality

Poor	Fair	Average	Good	Excellent

Sight Reading

Poor	Fair	Average	Good	Excellent

APPENDIX C

A Sample of Survey Letter
To Determine Weighted Value Scale

1805 Monroe
Commerce, Texas 75428
November 26, 1968

Dear Colleague:

I am currently working on a study of audition techniques. In the course of my study, I have become aware of the fact that most judges of individual student performances give more weight to some performance characteristics than they do others. This study is concerned with the type of audition which is conducted in selection of students for all-region and all-district bands.

In order for me to gain a more accurate insight into the proper weighting of performance characteristics to use in this study, I would like to ask you to indicate your personal percentage evaluation of the six characteristics listed below. Indicate the degree of emphasis you feel each of these qualities should receive in the spaces provided below.

1. Tone . . . . ______
2. Pitch. . . . . ______
3. Rhythm . . . . ______
4. Technical Accuracy . . ______
5. Musicality . . . . ______
6. Sight Reading. . . . . ______

100%

I realize how busy you are at this time, so I have tried to keep this request short. After you have listed these percentages, please return this letter to me in the stamped, self-addressed envelope which is enclosed. Thank you for your time.

Sincerely yours,

Curtis D. Owen, Jr.

CDO/bgo

Enclosure

APPENDIX D

Text of Instructions Given to Adjudicators Before Each Audition Session

TEXT OF INSTRUCTIONS GIVEN TO ADJUDICATORS BEFORE EACH AUDITION SESSION

Audition Session I

Gentlemen, thank you for your interest in this research project. The purpose of this audition session is to evaluate students who were recorded on tape when they auditioned for all-region band. Today you will use the first of three scoring techniques. Please take the scoring sheets out of the folder and examine the one on top which is labeled "sample." You will notice that six areas of performance have been listed. These are tone, pitch, rhythm, technical accuracy, musicality, and sight reading. Under each of these words is a continuum or a line which represents the poorest possible performance at the left end of the continuum and the best possible performance at the right end of the continuum. As each student is heard, you will determine and mark the point on the continuum you feel each student achieves with his overall performance.

There will be a pause before the sight reading material is played in order for you to mark the first five lines. The sixth line which is for sight reading is to be your evaluation of the total sight reading performance. In other words, the evaluation you make of sight reading will reflect the first five characteristics as they apply to sight reading. If you wish to make marks on the sheet while the student is playing, feel free to do so. However, let the mark which indicates the point of your evaluation be darker and heavier than any sub-marks you wish to use as a guide. Are there any questions at this point?

In your folders you will find a copy of the music which you will hear on the tape. The order of performance will be prepared work first, and then sight reading. In order for you to practice using this scoring technique, a sample recording will be played now. Use the scoring sheet marked "sample" for this practice. (A sample recording was played.)

You are to use one scoring sheet for each student. When you have finished scoring each student, turn the completed sheet face down on the desk in order to avoid using it twice. Do you have any questions?

The audition will now begin with student number one, scoring sheet number one.

Audition Session II

Gentlemen, today you will use the second scoring technique. If you recall, the first scoring sheet required that you evaluate six performance characteristics by marking a line through a continuum at the point you felt the student's performance had achieved. Open your folders and look at the yellow scoring sheet that you will use today for student number one. You will notice that the same performance characteristics are used in this scoring sheet as were used in the previous one. They are: tone, pitch, rhythm, technical accuracy, musicality, and sight reading. The big difference in this scoring sheet is that instead of numbers above the continuum to use as a guide, descriptions of these performance characteristics have been placed above each continuum. These descriptions picture the poorest possible performance on the left and the best possible performance on the right.

Look at the descriptions used for pitch. From left to right you will find: consistently out of tune, many serious deviations in pitch, several serious deviations in pitch, a few deviations in pitch, and very few or no deviations in pitch.

Notice in the descriptions which are more than one word long--that is the descriptions for tone, pitch, rhythm, and technical accuracy--many serious, several serious, a few, and a few or no,are terms used in each line to describe the characteristics.

You will mark these sheets with a dark line through the continuum at the point you feel the description describes the characteristic. Do you have any questions?

Since method number two of scoring is essentially the same as method number one, a sample will not be used. There will be a pause before sight reading in order for you to score the first five lines, a procedure which we did previously. Now review the descriptions briefly, and we will begin with student number one.

Audition Session III

Gentlemen, please open your folders and look at the scoring sheet for student number one. Today you will judge the all-region tapes using a procedure which will provide a standard of performance for you to use in your evaluation. You will hear a recording of a portion of the prepared work before each student is heard. This recording represents the description which you see at the right end of the continuum. Now look at your music. The major portion of the prepared work will be played now. (The entire recorded exercise was played by the professional at this point.)

Now that you have heard most of the prepared music, a short portion of this recording will be played before you hear each student perform. The scoring sheet you will use is the same as the one you used at our last judging session. There will be a pause before sight reading in order for you to score the first five lines. Do you have any questions?

A brief portion of the prepared work will be played first, and then each student will be heard. We are now ready to begin.

APPENDIX E

The Equal and Weighted Value Scales

EQUAL VALUE SCALE

Sight Reading	Musicality	Technical Accuracy	Rhythm	Pitch	Tone
0	0	0	0	0	0
.25	.25	.25	.25	.25	.25
.50	.50	.50	.50	.50	.50
.75	.75	.75	.75	.75	.75
1.00	1.00	1.00	1.00	1.00	1.00
1.25	1.25	1.25	1.25	1.25	1.25
1.50	1.50	1.50	1.50	1.50	1.50
1.75	1.75	1.75	1.75	1.75	1.75
2.00	2.00	2.00	2.00	2.00	2.00
2.25	2.25	2.25	2.25	2.25	2.25
2.50	2.50	2.50	2.50	2.50	2.50
2.75	2.75	2.75	2.75	2.75	2.75
3.00	3.00	3.00	3.00	3.00	3.00
3.25	3.25	3.25	3.25	3.25	3.25
3.50	3.50	3.50	3.50	3.50	3.50
3.75	3.75	3.75	3.75	3.75	3.75
4.00	4.00	4.00	4.00	4.00	4.00
4.25	4.25	4.25	4.25	4.25	4.25
4.50	4.50	4.50	4.50	4.50	4.50
4.75	4.75	4.75	4.75	4.75	4.75
5.00	5.00	5.00	5.00	5.00	5.00
5.25	5.25	5.25	5.25	5.25	5.25
5.50	5.50	5.50	5.50	5.50	5.50
5.75	5.75	5.75	5.75	5.75	5.75
6.00	6.00	6.00	6.00	6.00	6.00
6.25	6.25	6.25	6.25	6.25	6.25
6.50	6.50	6.50	6.50	6.50	6.50
6.75	6.75	6.75	6.75	6.75	6.75
7.00	7.00	7.00	7.00	7.00	7.00
7.25	7.25	7.25	7.25	7.25	7.25
7.50	7.50	7.50	7.50	7.50	7.50
7.75	7.75	7.75	7.75	7.75	7.75
8.00	8.00	8.00	8.00	8.00	8.00
8.25	8.25	8.25	8.25	8.25	8.25
8.50	8.50	8.50	8.50	8.50	8.50
8.75	8.75	8.75	8.75	8.75	8.75
9.00	9.00	9.00	9.00	9.00	9.00
9.25	9.25	9.25	9.25	9.25	9.25
9.50	9.50	9.50	9.50	9.50	9.50
9.75	9.75	9.75	9.75	9.75	9.75
10.00	10.00	10.00	10.00	10.00	10.00

WEIGHTED VALUE SCALE

Sight Reading	Musicality	Technical Accuracy	Rhythm	Pitch	Tone
0	0	0	0	0	0
.475	.600	.325	.300	.275	.525
.950	1.200	.650	.600	.550	1.050
1.425	1.800	.975	.900	.825	1.575
1.900	2.400	1.300	1.200	1.100	2.100
2.375	3.000	1.625	1.500	1.375	2.625
2.850	3.600	1.950	1.800	1.650	3.150
3.325	4.200	2.275	2.100	1.925	3.675
3.800	4.800	2.600	2.400	2.200	4.200
4.275	5.400	2.925	2.700	2.475	4.725
4.750	6.000	3.250	3.000	2.750	5.250
5.225	6.600	3.475	3.300	3.025	5.775
5.700	7.200	3.900	3.600	3.300	6.300
6.175	7,800	4.225	3.900	3.575	6.825
6.650	8.400	4.550	4.200	3.850	7.350
7,125	9.000	4.875	4.500	4.125	7.875
7.600	9.600	5.200	4.800	4.400	8.400
8.075	10.200	5.525	5.100	4.675	8.925
8.550	10.800	5.850	5.400	4.950	9.450
9.025	11.400	6.275	5.700	5.225	9.975
9.500	12,000	6.500	6.000	5.500	10,500
9.975	12.600	6.825	6.300	5.775	11,025
10.450	13.200	7.150	6.600	6.050	11.550
10.925	13.800	7.475	6.900	6.325	12.075
11.400	14.400	7.800	7.200	6.600	12.600
11.875	15.000	8.125	7.500	6.875	13.125
12.350	15.600	8.450	7.800	7.150	13.650
12.825	16.200	8.775	8.100	7.425	14.175
13.300	16.800	9.100	8.400	7.700	14.700
13.775	17.400	9.425	8.700	7.975	15.225
14.250	18.000	9.750	9.000	8.250	15.750
14.725	18.600	10.075	9.300	8.525	16.275
15.200	19.200	10.400	9.600	8.800	16.800
15.675	19.800	10.725	9.900	9.075	17.325
16.150	20.400	11.050	10.200	9.350	17.850
16.625	21.000	11.375	10.500	9.625	18.375
17.100	21.600	11.700	10.800	9.900	18.900
17.575	22.200	12.025	11.100	10.175	19.425
18.050	22.800	12.350	11.400	10.450	19.950
18.525	23.400	12.675	11.700	10.725	20.475
19.000	24.000	13.000	12.000	11.000	21.000

APPENDIX F

A List of the Weighted and Equal Value Rank Orders For the Four Instruments As Determined by the Twelve Adjudicators

RANK ORDER USING WEIGHTED VALUE SCALE FOR THE FLUTE AUDITION SESSIONS

Actual Ranking	Session I Judge			Session II Judge			Session III Judge		
	A	B	C	A	B	C	A	B	C
9.0	14.0	17.0	15.0	17.0	15.0	11.0	13.0	15.0	14.0
15.0	17.0	11.0	10.0	7.0	4.0	10.0	15.0	11.0	11.0
5.0	5.0	3.0	7.0	9.0	3.0	9.0	9.0	2.0	2.0
20.0	20.0	19.0	13.0	18.0	20.0	19.0	19.0	20.0	17.0
12.0	13.0	13.0	14.0	11.0	11.0	12.0	12.0	14.0	9.0
6.0	4.0	6.0	5.0	10.0	8.0	4.0	6.0	8.0	6.0
8.0	1.0	4.0	3.0	4.0	5.0	3.0	2.0	5.0	3.0
1.0	6.0	2.0	1.0	2.0	2.0	2.0	3.0	3.0	4.0
13.0	18.0	18.0	19.0	15.0	16.0	15.0	14.0	13.0	16.0
11.0	11.0	5.0	4.0	3.0	9.0	13.0	4.0	4.0	10.0
18.0	16.0	15.0	17.0	16.0	18.0	17.0	20.0	17.0	20.0
16.0	15.0	16.0	16.0	19.0	14.0	14.0	17.0	18.0	19.0
17.0	19.0	20.0	20.0	20.0	19.0	20.0	18.0	16.0	18.0
2.0	7.0	10.0	11.0	8.0	7.0	6.0	11.0	12.0	15.0
10.0	8.0	12.0	12.0	14.0	13.0	16.0	7.0	9.0	13.0
3.0	3.0	1.0	2.0	1.0	1.0	1.0	1.0	1.0	1.0
19.0	10.0	14.0	18.0	13.0	17.0	18.0	16.0	19.0	12.0
14.0	12.0	8.0	8.0	12.0	10.0	5.0	10.0	6.0	7.0
7.0	8.0	9.0	6.0	5.0	6.0	8.0	8.0	7.0	8.0
4.0	2.0	7.0	9.0	6.0	12.0	7.0	5.0	10.0	5.0

RANK ORDER USING EQUAL VALUE SCALE FOR THE FLUTE AUDITION SESSIONS

Actual Ranking	Session I Judge			Session II Judge			Session III Judge		
	A	B	C	A	B	C	A	B	C
9.0	14.0	17.0	15.0	17.0	15.0	11.0	13.0	15.0	13.0
15.0	17.0	11.0	10.0	8.0	3.5	10.0	15.0	11.0	11.0
5.0	5.5	3.0	8.0	9.0	3.5	9.0	9.0	2.0	2.0
20.0	20.0	20.0	14.0	18.0	20.0	19.0	19.0	20.0	17.0
12.0	13.0	13.0	13.0	10.0	11.0	12.0	11.5	14.0	9.0
6.0	4.0	6.0	5.0	12.0	7.5	4.0	6.0	8.0	6.0
8.0	1.0	4.0	2.0	4.0	5.0	3.0	2.0	5.0	3.0
1.0	5.5	2.0	1.0	2.0	2.0	2.0	3.0	3.0	4.0
13.0	18.0	18.0	18.0	15.0	16.0	15.0	14.0	13.0	16.0
11.0	10.5	5.0	4.0	3.0	9.0	13.0	4.0	4.0	10.0
18.0	16.0	15.0	17.0	16.0	17.0	17.0	20.0	17.0	20.0
16.0	15.0	16.0	16.0	19.0	14.0	14.0	16.5	18.0	18.5
17.0	19.0	19.0	20.0	20.0	19.0	20.0	18.0	16.0	18.5
2.0	7.0	10.0	11.0	7.0	7.5	7.0	10.0	9.0	15.0
10.0	8.0	12.0	12.0	13.5	12.5	16.0	7.0	12.0	14.0
3.0	3.0	1.0	3.0	1.0	1.0	1.0	1.0	1.0	1.0
19.0	10.5	14.0	19.0	13.5	18.0	18.0	16.5	19.0	12.0
14.0	12.0	8.0	7.0	11.0	10.0	5.0	11.5	6.0	7.0
7.0	9.0	8.0	6.0	5.0	6.0	6.0	8.0	7.0	8.0
4.0	2.0	8.0	9.0	6.0	12.5	8.0	5.0	10.0	5.0

RANK ORDER USING WEIGHTED VALUE SCALE FOR THE CLARINET AUDITION SESSIONS

	Session I			Session II			Session III		
	Judge			Judge			Judge		
Actual Ranking	A	B	C	A	B	C	A	B	C
10.0	12.0	9.0	10.0	4.0	7.0	10.0	6.0	3.0	5.0
13.0	19.0	12.0	20.0	15.0	14.0	17.0	14.0	12.0	13.0
15.0	11.0	8.0	12.0	19.0	11.0	9.0	13.0	17.0	15.0
20.0	7.0	18.0	6.0	14.0	16.0	19.0	20.0	19.0	16.0
5.0	3.0	1.0	5.0	2.0	3.0	6.0	2.0	4.0	3.0
12.0	15.0	15.0	15.5	16.0	12.0	16.0	19.0	14.0	17.0
16.0	18.0	16.0	17.0	18.0	18.0	20.0	16.0	18.0	18.0
3.0	10.0	5.0	8.0	8.0	10.0	8.0	9.0	10.0	11.0
11.0	17.0	14.0	18.0	13.0	6.0	12.0	12.0	11.0	9.0
8.0	13.0	19.0	15.5	11.0	5.0	14.0	4.0	7.0	12.0
9.0	8.0	2.0	9.0	9.0	13.0	3.0	5.0	6.0	8.0
17.0	14.0	20.0	13.0	10.0	15.0	11.0	8.0	13.0	10.0
7.0	5.0	3.0	7.0	7.0	8.0	5.0	11.0	9.0	7.0
6.0	6.0	7.0	4.0	5.0	9.0	7.0	10.0	8.0	6.0
19.0	9.0	13.0	14.0	17.0	20.0	13.0	18.0	16.0	19.0
2.0	1.0	4.0	3.0	1.0	2.0	1.0	7.0	2.0	4.0
1.0	2.0	6.0	2.0	3.0	1.0	4.0	3.0	1.0	2.0
18.0	20.0	17.0	19.0	20.0	19.0	18.0	15.0	15.0	20.0
4.0	4.0	10.0	1.0	6.0	4.0	2.0	1.0	5.0	1.0
14.0	16.0	11.0	11.0	12.0	17.0	15.0	17.0	20.0	14.0

RANK ORDER USING EQUAL VALUE SCALE FOR THE CLARINET AUDITION SESSIONS

Actual	Session I Judge			Session II Judge			Session III Judge		
Ranking	A	B	C	A	B	C	A	B	C
10.0	12.5	9.0	10.0	4.0	8.0	9.0	3.0	3.0	4.0
13.0	20.0	12.5	20.0	15.0	13.5	17.0	14.0	12.0	14.0
15.0	11.0	8.0	12.0	19.0	10.0	11.0	13.0	17.5	15.0
20.0	5.0	18.0	5.0	14.0	17.0	20.0	20.0	19.0	16.0
5.0	3.0	1.0	4.0	2.0	3.0	6.0	2.0	4.0	3.0
12.0	14.0	15.0	16.0	16.0	12.0	16.0	19.0	15.5	17.0
16.0	18.0	17.0	17.0	18.0	18.0	19.0	17.0	17.5	19.0
3.0	10.0	6.0	8.0	6.0	11.0	8.0	8.0	10.0	12.0
11.0	17.0	14.0	18.0	11.0	5.0	13.0	12.0	11.0	9.0
8.0	12.5	20.0	15.0	12.0	6.0	14.0	6.0	7.5	11.0
9.0	8.0	2.0	9.0	9.0	13.5	4.0	5.0	6.0	8.0
17.0	15.0	19.0	14.0	10.0	15.0	10.0	9.0	13.0	10.0
7.0	6.0	3.0	7.0	8.0	7.0	5.0	11.0	9.0	7.0
6.0	7.0	7.0	6.0	5.0	9.0	7.0	10.0	7.5	6.0
19.0	9.0	11.0	13.0	17.0	20.0	12.0	18.0	14.0	18.0
2.0	1.0	4.0	3.0	1.0	2.0	2.0	7.0	2.0	5.0
1.0	2.0	5.0	1.0	3.0	1.0	3.0	4.0	1.0	2.0
18.0	19.0	16.0	19.0	20.0	19.0	18.0	15.0	15.5	20.0
4.0	6.0	10.0	2.0	7.0	4.0	1.0	1.0	5.0	1.0
14.0	16.0	12.5	11.0	13.0	16.0	15.0	16.0	20.0	13.0

RANK ORDER USING WEIGHTED VALUE SCALE FOR THE CORNET-TRUMPET AUDITION SESSIONS

Actual Ranking	Session I Judge			Session II Judge			Session III Judge		
	A	B	C	A	B	C	A	B	C
19.0	16.0	10.0	15.0	11.0	6.0	12.0	4.0	8.0	9.0
4.0	9.0	7.0	9.0	7.0	11.0	11.0	15.0	3.0	12.0
11.0	20.0	8.0	8.0	14.0	9.0	14.0	14.0	15.0	11.0
15.0	13.0	19.0	12.0	20.0	17.0	17.0	18.0	18.0	16.0
1.0	2.0	9.0	7.0	4.0	3.5	9.0	1.0	4.0	4.0
20.0	19.0	17.0	17.0	17.0	19.0	19.0	19.0	16.0	19.0
13.0	4.0	4.0	10.0	9.0	1.0	5.0	5.0	5.0	2.0
9.0	6.0	3.0	6.0	1.0	7.0	6.0	3.0	10.0	1.0
16.0	10.0	11.0	14.0	12.0	10.0	7.0	11.0	12.0	14.0
18.0	18.0	20.0	20.0	18.0	16.0	16.0	16.0	20.0	20.0
8.0	12.0	15.0	16.0	13.0	15.0	10.0	2.0	13.0	10.0
3.0	1.0	2.0	2.0	2.0	8.0	3.0	12.0	6.0	6.0
10.0	7.0	12.0	11.0	6.0	13.0	4.0	9.0	14.0	7.0
14.0	17.0	18.0	18.0	19.0	18.0	18.0	20.0	19.0	17.0
7.0	5.0	5.0	5.0	5.0	2.0	1.0	8.0	1.0	3.0
17.0	14.0	16.0	19.0	15.0	20.0	20.0	17.0	17.0	18.0
12.0	8.0	14.0	13.0	16.0	12.0	15.0	6.0	9.0	15.0
2.0	3.0	1.0	4.0	3.0	5.0	2.0	10.0	2.0	8.0
6.0	15.0	13.0	3.0	10.0	14.0	13.0	13.0	7.0	5.0
5.0	11.0	6.0	1.0	8.0	3.5	8.0	7.0	11.0	13.0

RANK ORDER USING EQUAL VALUE SCALE FOR THE CORNET-TRUMPET AUDITION SESSIONS

Actual Ranking	Session I Judge			Session II Judge			Session III Judge		
	A	B	C	A	B	C	A	B	C
19.0	16.0	11.0	15.0	12.0	6.0	13.0	2.0	11.0	9.0
4.0	9.0	7.5	9.0	7.0	11.0	11.0	14.0	3.0	11.0
11.0	20.0	7.5	7.0	16.0	7.0	14.0	13.0	15.0	13.0
15.0	13.0	19.0	11.0	20.0	17.0	16.0	18.5	18.0	16.0
1.0	2.0	9.0	6.0	4.0	3.5	9.0	1.0	4.0	4.0
20.0	17.0	18.0	17.0	17.0	20.0	20.0	20.0	17.0	18.0
13.0	4.0	5.0	10.0	9.0	1.0	6.0	5.0	5.0	2.0
9.0	6.0	4.0	8.0	1.0	9.5	5.0	4.0	9.0	1.0
16.0	10.0	10.0	14.0	11.0	9.5	8.0	12.0	11.0	15.0
18.0	18.5	20.0	20.0	18.0	16.0	17.0	17.0	20.0	19.0
8.0	12.0	15.0	16.0	13.0	15.0	10.0	3.0	13.0	10.0
3.0	1.0	2.0	2.0	2.0	8.0	4.0	11.0	7.0	7.0
10.0	7.0	12.5	12.0	6.0	14.0	3.0	9.5	14.0	8.0
14.0	18.5	17.0	18.5	19.0	18.0	18.0	18.5	19.0	17.0
7.0	5.0	3.0	5.0	5.0	2.0	1.0	8.0	1.5	3.0
17.0	14.0	16.0	18.5	14.0	19.0	19.0	15.0	16.0	20.0
12.0	8.0	14.0	13.0	15.0	12.0	15.0	6.0	8.0	14.0
2.0	3.0	1.0	4.0	3.0	5.0	2.0	9.5	1.5	5.5
6.0	15.0	12.5	3.0	10.0	13.0	12.0	16.0	6.0	5.5
5.0	11.0	6.0	1.0	8.0	3.5	7.0	7.0	11.0	12.0

RANK ORDER USING WEIGHTED VALUE SCALE FOR THE TROMBONE AUDITION SESSIONS

Actual Ranking	Session I Judge			Session II Judge			Session III Judge		
	A	B	C	A	B	C	A	B	C
5.0	7.0	3.0	3.0	4.0	4.0	5.0	3.0	7.0	4.0
8.0	6.0	6.0	6.0	8.0	13.0	9.0	9.0	14.0	6.0
14.0	10.0	14.0	4.0	14.0	11.0	12.0	10.0	6.0	14.0
16.0	16.0	10.0	15.0	15.0	16.0	15.0	15.0	17.0	15.0
2.0	12.0	7.0	11.0	3.0	5.0	7.0	6.0	2.0	5.0
15.0	17.0	9.0	16.0	16.0	15.0	14.0	14.0	12.0	16.0
11.0	13.0	13.0	13.0	9.0	12.0	10.0	13.0	8.0	12.0
13.0	11.0	15.0	14.0	13.0	14.0	16.0	16.0	15.0	17.0
9.0	9.0	16.0	10.0	11.0	8.0	11.0	12.0	11.0	11.0
3.0	1.0	2.0	2.0	2.0	2.0	3.0	7.0	3.0	2.0
4.0	3.0	12.0	7.0	5.0	3.0	4.0	1.0	4.0	7.0
17.0	15.0	17.0	17.0	17.0	17.0	17.0	17.0	16.0	13.0
7.0	5.0	11.0	8.0	10.0	7.0	8.0	8.0	5.0	10.0
12.0	8.0	5.0	9.0	7.0	9.0	13.0	5.0	13.0	9.0
1.0	2.0	1.0	1.0	1.0	1.0	1.0	2.0	1.0	1.0
6.0	4.0	4.0	5.0	6.0	6.0	2.0	4.0	9.0	3.0
10.0	14.0	8.0	12.0	12.0	10.0	6.0	11.0	10.0	8.0

RANK ORDER USING EQUAL VALUE SCALE FOR THE TROMBONE AUDITION SESSIONS

Actual Ranking	Session I Judge			Session II Judge			Session III Judge		
	A	B	C	A	B	C	A	B	C
5.0	7.0	3.0	3.0	4.0	5.0	5.0	3.0	7.0	4.0
8.0	6.0	6.0	6.0	8.0	14.0	9.0	9.0	14.0	6.0
14.0	9.0	14.0	4.0	14.0	11.0	13.0	10.0	6.0	14.0
16.0	16.0	10.0	15.0	15.0	15.0	15.0	15.0	17.0	15.0
2.0	10.0	7.0	11.0	2.5	4.0	7.0	6.0	3.0	5.0
15.0	17.0	9.0	16.0	16.0	16.0	14.0	14.0	12.0	16.0
11.0	13.0	13.0	13.5	9.0	12.0	10.0	13.0	8.0	12.0
13.0	11.0	15.0	13.5	12.0	13.0	16.0	16.0	15.0	17.0
9.0	12.0	16.0	10.0	11.0	9.0	11.0	12.0	11.0	11.0
3.0	1.0	2.0	2.0	2.5	2.0	3.0	7.0	4.0	2.0
4.0	3.0	12.0	8.0	5.0	3.0	4.0	1.0	2.0	7.0
17.0	15.0	17.0	17.0	17.0	17.0	17.0	17.0	16.0	13.0
7.0	5.0	11.0	7.0	10.0	7.0	8.0	8.0	5.0	10.0
12.0	8.0	5.0	9.0	7.0	8.0	12.0	5.0	13.0	9.0
1.0	2.0	1.0	1.0	1.0	1.0	1.0	2.0	1.0	1.0
6.0	4.0	4.0	5.0	6.0	6.0	2.0	4.0	9.0	3.0
10.0	14.0	8.0	12.0	13.0	10.0	6.0	11.0	10.0	8.0